

POLYADENYLATION OF RIBOSOMAL RNA IN RESPONSE TO PICORNAVIRUS  
INFECTION IN HONEY BEES (*APIS MELLIFERA*)

BY

JOHNNY YU

THESIS

Submitted in partial fulfillment of the requirements  
for the degree of Master of Science in Entomology  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Advisers:

Professor Matthew E. Hudson  
Professor Gene E. Robinson  
Professor May R. Berenbaum

# Abstract

At least eighteen different picornaviruses can infect honey bees (*Apis mellifera*). The combination of the virulence of these pathogens as well as their transmissibility through *Varroa destructor*, an ectoparasitic mite, can cause severe declines in honey bee populations.

Polyadenylated (polyA) rRNA has been associated with honey bees infected with viruses in past studies. In bacteria, yeast, and some mammals, polyA rRNA serves as a molecular marker for the degradation of rRNA. Honey bees may be utilizing polyA rRNA in the same fashion to degrade cells that are infected with picornaviruses. In this thesis, I attempted to correlate viral infection and polyA rRNA using qRT-PCR and RNA-seq. Five polyA regions were identified with RNA-seq. One region (28S 2070-2130) was weakly associated with viral infection. I used quantitative RT-PCR to more accurately measure the 28S 2070-2130 region and again found a weak association between viral infection and polyA rRNA. Also, significantly more polyA rRNA was found in the abdomen than in the head. The high expression levels of polyA rRNA in the abdomen may be a consequence of picornaviruses that infect mostly the abdomen.

# Acknowledgments

I would like to thank my advisors, Matthew Hudson, Gene Robinson, and May Berenbaum, for their guidance in performing and finishing this thesis. Half of this project could not have happened without Radhika Khetani, who not only provided me with the RNA-seq data, but also guided me through the process of mining for polyA reads and performed the DEGseq analysis. Samples from the study were given by Danielle Downey, the Apiculture Specialist at the Hawaii Department of Agriculture, and Chester Danbury of Danbury Apiaries in Kilauea, Hawaii. They meticulously procured the *Varroa*-free honey bees in a liquid nitrogen Dewar flask and shipped it to Illinois. I owe a debt of gratitude to Thomas Newman for his help in the shipment of bees and designing primers. Kranthi Vrala wrote Perl scripts that trimmed and modified the Illumina data. The Perl script used to find SNPs was written by Indrajit Kumar. Finally, I would like to thank Tania Jogesh for all her love and support in the completion of this thesis.

# Table of Contents

1. Introduction .....	1
2. Literature Review .....	6
2.1 General honey bee biology .....	6
2.2 Colony Collapse Disorder .....	7
2.3 Honey bee viruses .....	9
2.4 Insect immunity.....	11
2.5 Molecular techniques for pathogen analysis.....	13
3. Materials and Methods.....	17
3.1 Comparison of Hawaiian virus-low bees and USDA virus-high bees using qRT-PCR .....	17
3.2 Mining for polyA rRNA sequences in RNA-seq data .....	20
3.3 PCR of polyA regions discovered in RNA-seq analysis .....	23
4. Results.....	25
5. Discussion.....	28
5.1 Initial qRT-PCR.....	28
5.2 RNA-seq analysis .....	29
5.3 Conclusion .....	33
6. Tables and Figures.....	36
References .....	57
Appendix A: Shell script to find polyA sequences.....	69
<b>Appendix B: Perl script to find SNPs .....</b>	<b>71</b>

# 1. Introduction

The European honey bee, *Apis mellifera* L., is an indispensable economic asset to the agricultural industry worldwide (Aizen et al. 2008; Gallai et al. 2009). The honey bee is responsible for approximately \$15 billion worth of pollination services annually (Calderone, 2012). Since 2006, honey bee losses have been increasing, especially in the United States, where some apiculturists report decreases as high as ninety percent (Ellis et al., 2010). Apiculturists and scientists have termed this dramatic loss in honey bees as Colony Collapse Disorder (CCD), a syndrome where the workers inexplicably abandon the colony and never return, leaving the colony to “collapse” (vanEngelsdorp et al. 2009). Scientists have not been able to find definitive causes for CCD and now the scientific community concurs that a suite of pathogens, parasites, and other stress factors may be responsible (vanEngelsdorp et al. 2009).

*A. mellifera* is challenged by a variety of pathogens and pests, including bacteria, fungi, and other invertebrates. Invasive species that originally parasitized the Asian honey bee, *A. ceranae* Fabr., have become even more destructive after they shifted hosts to *A. mellifera* colonies. One such invasive pest, *Varroa destructor* Anderson & Trueman, an ectoparasitic mite, has been suspected to be a contributing factor of CCD (Highfield et al., 2009; Rosenkranz et al. 2010). Mites infest a colony and feed on hemolymph from honey bee pupae and adults. This parasitic interaction not only drains the honey bee of bodily fluids, but also transmits pathogens (Chen et al. 2006). During the feeding process, the mite vectors many honey bee viruses belonging to a group of single stranded RNA (ssRNA) viruses called picornaviruses. At

least eighteen viruses in this family have been found to cause sickness in the honey bee, the most ubiquitous being Deformed Wing Virus (DWV) (Bowen-Walker et al. 1999). DWV has had a sublethal impact on honey bees in the past, but now in conjunction with *V. destructor*, DWV and other picornaviruses have even more of a deleterious impact on honey bee health and together serve as an indicator for CCD (Chen et al. 2005; Dainat et al. 2012). *V. destructor* and DWV appeared in the western hemisphere years before the first occurrence of CCD (Nordström, 2003), but only recently have been associated with CCD. The presence of the mite and virus before and after CCD insinuates that they cannot be the only factors in CCD. Nonetheless they have a significant negative impact on honey bee health.

The *Varroa* mite also suppresses the honey bee immune system, which leads to increased viral replication (Amdam et al., 2004; Gregory et al., 2005; Yang & Cox-Foster, 2005). Insect cellular and humoral immune responses, including antimicrobial peptides and enzymes involved in degrading foreign substances, are down-regulated in honey bees infested with the *Varroa* mite. The down-regulation of these invertebrate immune responses facilitates infection by picornaviruses. The reduction of general cellular and humoral responses has been linked to viral amplification via the fact that the honey bee becomes so immunocompromised due to unchecked invasive bacteria and toxins that viruses have the opportunity to rapidly multiply (Shen et al. 2005).

*A. mellifera* live in a very dense population that is prone to pathogen transmission (Schmid-Hempel, 2005). Even so, in a genome-wide comparison between *A. mellifera*, *Drosophila melanogaster* Meig., and *Anopheles gambiae* Giles, *A. mellifera* had the fewest immunity-related genes (Evans et al. 2006). However, honey bees may possess unidentified

immune responses that were not accounted for in this comparison of known immunity-related genes. For example, a unique immune responsive protein, IRP30, has only been observed in hymenopterans, but no other insects (Albert et al., 2011). The immune response of *A. mellifera* to viruses is still unknown (Azzami et al., 2012). Honey bees have evolved behavioral hygienic traits to defend themselves against diseases, so the evolution of a novel cellular response to viral infection unique to *A. mellifera* may also be possible (Evans et al. 2006).

*A. mellifera* can survive with a chronic infection of DWV, so a defense may exist to protect *A. mellifera* from an acute infection of picornaviruses. In this project I investigated an alternative immune response that has not yet been confirmed in *A. mellifera*: ribosomal RNA (rRNA) polyadenylation. Polyadenylated (polyA) rRNA is unusual because, in most cases, polyadenylation occurs at the ends of messenger RNA. Johnson et al. (2009) found that accumulation of polyA rRNA was associated with infection by multiple picornavirus species and with diagnosis of CCD. Cells that experience stress can create degradation intermediate for cellular apoptosis by rRNA polyadenylation, as seen in bacteria, yeast, mice and humans (Grzechnik & Kufel, 2008; Kong et al., 2008; Lange et al., 2008; Slomovic et al., 2006a; Slomovic et al., 2006b). Cell lines of the fall armyworm, *Spodoptera frugiperda* Smith, exhibits apoptotic responses when inoculated with baculovirus (Clarke & Clem, 2003). Polyadenylation leading to cellular apoptosis would be a novel defense against picornaviruses in *A. mellifera*.

To test for ribosomal polyadenylation as a response to infection by picornaviruses, two groups of honey bees were collected. One group of bees acted a control and originated from Kauai, Hawaii, a location where the *Varroa* mite has not yet reached (Martin et al., 2012). The absence of the major vector reduced the likelihood that Hawaiian honey bees would have a

high picornavirus infection. If this polyA rRNA is indeed associated with viral infection, the Hawaiian bees would be expected to display little to no polyA rRNA. The second group of bees, previously diagnosed to be with an infection with at least one picornavirus, acted as the experimental group and originated from USDA labs in Beltsville, MD. I hypothesized that the experimental group of bees to have a higher expression level of polyA rRNA. I used qRT-PCR to measure the quantity of polyA rRNA and two picornaviruses, Deformed Wing Virus (DWV) and Black Queen Cell Virus (BQCV).

Ribosomal polyadenylation sites were also mined for in RNA-seq libraries that were opportunistically obtained from two past *A. mellifera* gene expression experiments performed by Khetani & Evans (unpublished data, 2012) and Alaux et al. (2009). Even though these studies were not specifically designed to test changes in gene expression in response to viral infection, most honey bee colonies in the United States are infested with *Varroa*, and in turn have an unapparent chronic infection of picornavirus, so the hypothesized polyadenylation reaction may exist in bees that were not specifically diagnosed for a picornavirus (Tentcheva et al., 2004). Other honey bee pathogens present in these experiments may affect comparisons between virus-low and virus-high bees. Nonetheless, it would be difficult to obtain samples with just one type of infection as honey pathogens are so widespread. Most honey bee samples that could be obtained would most likely be co-infected with another pathogen (Chen et al., 2004).

I also attempted to find an association between viral load and polyA rRNA quantity in the two RNA-seq libraries obtained from Khetani & Evans (unpublished data, 2012) and Alaux et al. (2009). However, in Johnson et al. (2009), higher expression levels of polyA rRNA was found

to be associated with infection of multiple picornavirus species, rather than increasing viral load. I only measured polyA rRNA in response to viral load because insects have a non-specific immune response that may not respond uniquely based on the species of viruses (Schmid-Hempel, 2005). Instead of virus diversity, I expect viral load to play a role in the initiation of polyA rRNA as an immune response.

## 2. Literature Review

### 2.1 General honey bee biology

*A. mellifera*, as reviewed in Winston (1991), is eusocial insect that lives in large colonies with 20,000 to 60,000 individuals. A single queen lays eggs and is responsible for all of the reproduction in the hive. Before the queen establishes a colony, she participates in a mating swarm when she can mate with up to 20 male honey bees (drones). The queen uses the sperm collected from this mating flight to produce tens of thousands of bees for the rest of her life. Honey bees taxonomically belong in the order Hymenoptera and, like all hymenopterans, are haplodiploid. Every fertilized egg becomes a female and every unfertilized egg becomes a drone. In honey bees, the daughters of the queen will generally not reproduce but instead only help maintain the colony and are called workers. Workers in the colony do not typically produce young, but instead are responsible for performing other hive tasks. Workers nurse the young, maintain the hive, forage for water and food, and guard the hive.

Honey bees forage for nectar in numerous different types of flowering plants to convert to honey to feed the colony. While honey bees forage for nectar and pollen, pollen can adhere to the bodies of the bees and be transferred to other conspecific plants in a process called pollination. Pollination assists the sexual reproduction of many flowering plants, leading to plant populations with greater genetic diversity, a necessity for ecosystem health. Many cultivated plants, notably most fruit and nut trees, cannot disperse their pollen or produce high crop yields without honey bee pollination. Pollination services provided by honey bees and

other insect pollinators account for approximately 35% of food consumed by people and produce \$15 billion worth of food annually (Calderone, 2012; Gallai et al., 2009).

## **2.2 Colony Collapse Disorder**

In the fall of 2006, managed honey bee colonies declined by 30-90% (Ellis et al., 2010). Beekeepers observed that the collapsing hives were healthy until the foragers abruptly abandoned the colony, leaving behind brood, honey stores, and the queen (Oldroyd, 2007). Apiculturists and biologists termed this combination of symptoms as Colony Collapse Disorder (CCD). Years after the first observations of CCD, beekeepers worldwide would continue to diagnose their hives as suffering from CCD to this date (Dainat et al. 2012; vanEngelsdorp et al. 2011).

Scientists have tried to identify recent changes in apiculture that could have caused CCD. Of these changes in apiculture, many honey bee enthusiasts have identified chemical pesticides as a possible cause of CCD. Honey bees have an exceptionally low number of detoxification enzymes, which in theory makes them even more susceptible to pesticides than other insects (Claudianos et al., 2006). Pesticides used in crop production have been implicated in decreasing forager survival, a trait similar to the symptoms of CCD (Henry et al., 2012). Pesticide residues found in the pollen collected by honey bees indicate that these harmful chemicals can be ingested and support the contention that pesticides can do serious harm to honey bees (Chauzat et al., 2006). However, not all colonies that succumb to CCD face the same pesticide or suite of pesticides, so pesticides have not been determined as the definitive cause of CCD (Mullin et al., 2010).

In addition to chemical dangers, a variety of biological pests and pathogens threaten honey bees. *Nosema apis* Z. and *N. ceranae*, two species in the group microsporidia, have caused major declines in honey bee populations in the past ten years (Martín-Hernández et al., 2007). Fungal spores from these microsporidians infect the midgut of honey bees and multiply in their bodies, eventually leading to organ failure and death (Higes et al., 2007). In the 1990s, *N. apis* was the most prevalent microsporidian to infect bees in the Western Hemisphere (Klee et al., 2007). Later in the 2000s, *N. ceranae* became more prevalent, even though the fungus had already been present in the Western Hemisphere for the past twenty years (Chen et al., 2008). Some reports claim that incidence of *N. ceranae* is associated with CCD, but its presence in both healthy and CCD hives indicates it is not the only factor in CCD (Oldroyd, 2007).

Other parasites have been suspected to cause CCD. *Apocephalus borealis* Brues, a phorid fly, is a parasitoid to honey bees and can alter their behavior such that infested bees actually abandon the hive and die after the fly emerges from their body (Core et al., 2012). The sudden abandonment of a hive is similar to the behavior seen in CCD bees, but the phorid fly does not occur in honey bee colonies often enough to explain the collapse of so many hives in 2006.

Mites, especially *V. destructor*, have played a large role in honey bee colony decline for the past thirty years in American apiculture (Ball & Allen, 1988). Honey bee tracheal mites, *Acarapis woodi* Rennie, and the Asian parasitic brood mite, *Tropilaelaps clareae* Del. And Baker, have the potential to infest honey bees and can cause death, but *V. destructor* has been more prevalent and destructive to colonies. *V. destructor* originated in Asia and spread to North America, Europe, Africa and even isolated islands, such as Hawaii and New Zealand (Martin et al., 2012; Zhang, 2000). *V. destructor* latches onto honey bees and feeds by draining the hemolymph from

pupae and adults. Asian honey bees (*A. cerana*) and Africanized honey bees (*A. mellifera scutellata* Hepburn and Radloff) groom colony members more diligently, but European honey bees (EHB) are less hygienic and are more susceptible to mites and other parasites (Spivak & Reuter, 2001). *V. destructor* can be controlled with miticides, but the over-application of these miticides can decrease the health of the colony, so mites still persist in apiaries (Pettis, 2004). The persistent infestation of hemolymph-draining mites has been consistently ranked as one of the contributing factors of CCD (Dainat et al., 2012).

### **2.3 Honey bee viruses**

The spread of *V. destructor* through migratory beekeeping has increased the incidence and virulence of deleterious honey bee viruses (Sumpter & Martin, 2004). Viral pathogens have been less of a concern to beekeepers relative to the honey bee threats mentioned above, but recently, viruses in combination with the *Varroa* mite have been so deadly that they even have been implicated as a cause for CCD (Bromenshenk et al., 2010; Cox-Foster et al., 2007). Honey bee viral infections typically are latent as well as unapparent, meaning that bees are symptomless even though they possess the virus, and only become acute and lethal if the bee becomes immunocompromised (Ribi re et al., 2008).

At least eighteen different viruses infect the honey bee, the majority of which are picornaviruses, positive strand single stranded RNA viruses. The most prevalent of these are Deformed Wing Virus (DWV), Black Queen Cell Virus (BQCV), Sacbrood Virus (SBV), Kashmir Bee Virus (KBV), Acute bee paralysis virus (ABPV), and Chronic Bee Paralysis Virus (CBPV) (Chen & Siede, 2007; vanEngelsdorp et al., 2009). They can infect the larval, pupal and adult stages of

the honey bee (Aubert, 2008). Multiple viral infections can be present in a single bee (Chen et al., 2004). BQCV and DWV are the most prevalent because they can be vertically transmitted through an infected queen that is responsible for the reproduction of an entire colony (Chen et al., 2006).

BQCV gives rise to dead queen larvae in black cells (Bailey & Woods, 1977). Many nurse bees carry a low level of BQCV. Several nurses that together feed a queen larva transmit a high enough quantity of BQCV that the queen becomes heavily infected. Queens either die from infection at this stage or mature into an adult queen. The queen can live on, not showing any sign of infection, until the bees she mothers eventually nurse a future queen (Nordström et al., 1999). A mature infected adult queen carries the virus in her gut as well as her ovaries, meaning that she can vertically transmit the virus to her offspring (Chen et al., 2006). Like the other honey bee viruses, BQCV can be transmitted fecally, orally or through the vector *V. destructor* (Chantawannakul et al., 2006).

The most common and virulent of all the picornaviruses that infect honey bees is DWV (Miranda & Genersch, 2010). DWV is most apparent in adult honey bees, but can infect all life stages of the honey bee. The symptoms are most noticeable in adult bees that have “shrunk, crumpled wings, decreased body size, and discoloration” and a decreased lifespan (Chen & Siede, 2007). DWV was once considered to be a virus of low pathogenicity because of its persistence in colonies at high titers, but in combination with parasitization by *V. destructor*, DWV has become more virulent (Bowen-Walker et al. 1999; Gisder et al., 2009; Shen et al., 2005).

All viruses use the host ribosome to synthesize their own proteins, but picornaviruses can synthesize the necessary viral proteins with fewer steps than some other viruses, as they consist of a single-stranded RNA which is sense (+) strand and thus immediately capable of translation once within the cell. In picornavirus genomic RNA, the 5' UTR contains the Internal Ribosomal Entry Site (IRES) that allows the virus to enter the ribosome more quickly than the host mRNA and synthesize its proteins. Usually ribosomes recognize the 5' methylated cap of mRNA transcripts and subsequently use eukaryotic initiation factors (EIFs) to carry out translation (Martínez-Salas et al., 2001). Picornaviruses, on the other hand, use the IRES to gain entry into the ribosome, need fewer EIFs to start translation, and even cleave EIFs needed for host mRNA translation (Pestova & Hellen, 2006). To this date, no conserved IRES sequence has been established, but the IRES sequence is assumed to undergo high mutation rates as with the rest of the viral genome (Fernández-Miragall & Martínez-Salas, 2007). Part of their virulence may be attributed to the RNA virus's high mutation rate (Drake et al., 1998). Picornaviruses use an RNA-dependent RNA polymerase to replicate their genome. RNA-dependent RNA polymerases are more error-prone than DNA-dependent DNA polymerases, which most organisms use to replicate their genome. The high mutation rate due to the lack of proofreading activity by RNA polymerase lends RNA viruses the ability to adapt to host responses.

## **2.4 Insect immunity**

Insects lack the adaptive immune system seen in vertebrates. After the invasion of a foreign microbe or virus, insects respond with innate immune responses, not an adaptive

response that uses the memory of a past attack. The insect innate immune system has two responses: the cellular response and the humoral response. The cellular response consists of phagocytes that circulate in the hemolymph and engulf invading bacteria and fungi (Strand, 2008). The humoral response consists of the Toll and immune deficiency (Imd) pathways, which activate antimicrobial peptides (AMPs) that also attack bacteria and fungi (De Gregorio et al., 2002).

To this date, most immune responses studied in *A. mellifera* have been in relation to bacteria and fungi. The immune response to viruses remains an active field of study. Research in the field of insect virology has been limited by the lack of understanding mechanisms behind the defenses observed in insects. For example, the activation of the Toll pathway is associated with viral defense in *D. melanogaster*, but the mechanism by which the Toll pathway uses AMPs to attack virus is unknown (Zamboni et al., 2005). In a similar case, NF- $\kappa$ B, a transcription factor in the Imd pathway, leads to upregulation of AMPs and other antiviral responses associated with viral infection, but the molecular mechanism of the response is again unknown (Avadhanula et al., 2009). NF- $\kappa$ B was recently found to be downregulated in *Varroa*-infested honey bees, but again, the physiological consequences of the downregulation has yet to be determined (Azzami et al., 2012; Nazzi et al., 2012). When inoculated with a picornavirus, *Drosophila C Virus* (DCV), *D. melanogaster* induces the transcription of the *vir-1* gene through the Jak-STAT pathway (Dostert et al., 2005). While *vir-1* is upregulated in response to DCV, the mode of action of *vir-1* has not been characterized in association with an actual physiological response.

The only well-studied viral response mechanism is encapsulation in lepidoterans and RNA interference (RNAi) in dipterans. In lepidopterans, the cellular immune system responds to baculoviruses by encapsulation using hemocytes that circulate in the hemolymph (Washburn et al., 1996). Unlike lepidopterans, honey bees face mostly ssRNA viruses, not double-stranded DNA (dsDNA) baculoviruses, and may elicit other undiscovered cellular responses. Mosquitoes and fruit flies utilize RNAi to break down RNA viruses by cleaving the dsRNA form of RNA viruses with Dicer, an endoribonuclease, and then attaching the product to the RNA-Induced Silencing Complex (RISC) for degradation (Arjona et al., 2011; Wang et al., 2006). To this date, honey bees have been observed to utilize RNAi only in response to Israeli Acute Paralysis Virus (IAPV), another picornavirus (Hunter, 2010). The RNAi mechanism is present in most organisms as a mechanism to manage gene regulation and antiviral response (Hannon, 2002).

## **2.5 Molecular techniques for pathogen analysis**

Molecular techniques exist to identify the RNA in present in a honey bee sample. Serological and histological techniques can detect viruses, but molecular techniques are more quantitative, sensitive, accurate, and cheaper to perform (Miranda, 2008). RNA extractions from a honey bee can give information about the bee itself and also the pathogens infecting the bee. To identify and measure the types of RNA present in a honey bee, the RNA must first be reverse-transcribed in to cDNA, a more stable molecule that can be analyzed in downstream processes such as PCR, microarrays, and DNA sequencing, namely high-throughput sequencing (HTS). In quantitative reverse transcription PCR (qRT-PCR), cDNA is used to measure specific genes that can be targeted with primers. A fluorescent dye binds to the double-stranded DNA

that is detected by the instrument. The amount of fluorescence in a sample reflects how much PCR product, or gene of interest, is present in the sample. This method allows for accurate quantification of honey bee genes as well as viruses present in a sample (Miranda, 2008).

Pragmatically speaking, only a few out of several thousands of genes can be analyzed with qRT-PCR. Microarrays and HTS on the other hand can analyze thousands of genes in a single reaction. Microarrays, HTS, and other similar biological technologies produced so much data that new questions could be asked and new data analysis techniques were created, all of which led to the creation of a new field of biology, genomics. Genomics is divided into subcategories based on the type of molecule being compared. In transcriptomics, gene regulation is compared by converting mRNA into cDNA and then identifying and quantifying transcript levels between groups. In metagenomics, all of the RNA or DNA in a sample is sequenced to find all of the species of microbiota and viruses (Hudson, 2008). Microarrays are a widely used technology which measures the expression of genes in an organism by putting known gene sequences onto a “chip” or other microscopic DNA detection device to which cDNA can bind (Schena et al., 1995). In recent years, researchers have attempted to quantify honey bee responses to challenges such as *Varroa* infestation and CCD by looking at gene expression using microarrays (Navajas et al., 2008; Johnson et al., 2009). Typically, a microarray is used to test gene regulation, but the chip can also be used to identify and measure pathogens in a sample (Johnson et al., 2009; Runckel et al., 2011).

Microarrays are an affordable and reliable way to measure the transcription of many genes present in a sample, but the field of genomics has begun transitioning toward HTS and RNA-seq analyses (Marioni et al., 2008). Instead of amplifying one region as in qRT-PCR or

relying on probes, HTS employs massive parallel sequencing to identify every base within cDNA libraries. Past studies that have done the same experiment with both RNA-seq and microarrays have demonstrated that the data generated from RNA-seq revealed more differentially-expressed genes and better quantified differences in gene expression (Mortazavi et al. 2008; Nagalakshmi et al. 2008). RNA-seq has a number of advantages over microarrays. RNA-seq is a more robust and sensitive method that detects transcripts in low levels that would not be detected by microarrays (Wold & Myers, 2008). Unlike RNA-seq, microarrays frequently produce false positives because of cross-hybridization (Weinstein et al., 2002). Microarray studies are also difficult to reproduce because of the variability in microarray production and data analysis (Draghici et al., 2006). RNA-seq also removes the bias of *a priori* gene selection, as microarrays can only measure the expression of genes and pathogens with sequences that are known in advance, which narrows the potential of the study by excluding the possibility of discovering new genes and pathogens. RNA-seq, however, gives the sequence and abundance of the transcripts of all genes whether or not they are known in advance. As sequencing technologies become cheaper because of technological advances, the cost of performing RNA-seq is falling relative to utilizing a microarray (Wold & Myers, 2008).

Metagenomics analyses have revealed viruses never detected before in honey bees, but the results have not revealed the reason behind CCD. For example, IAPV was discovered to occur more frequently in bees diagnosed with CCD than in healthy bees by using metagenomic analyses (Cox-Foster et al., 2007). The pathogen was first believed to be the new, introduced factor that caused CCD, but IAPV was later on discovered to have been present in American honey bee populations for decades and present in both healthy and declining hives (Chen &

Evans, 2007). Also, the results from these new molecular techniques can be misinterpreted. These new molecular techniques rely on measuring molecular/proteomic information instead of phenotypic infection in live honey bees, so claims of an association of a particular pathogen with CCD are sometimes exaggerated. For instance, invertebrate iridescent virus-6 (IIV-6) was thought to be linked to CCD in a study that used mass spectrometry-based proteomics to detect proteins present in healthy bees and CCD bees (Bromenshenk et al., 2010). This association was later discredited as the authors failed to include *A. mellifera* proteins in their analysis and falsely identified proteins as IIV-6 when the proteins actually had more homology with *A. mellifera* proteins (Foster, 2011). Sometimes new, possibly pathogenic organisms are found in *A. mellifera* but have no known pathogenic effects. Lake Sinai Virus (LSV) and *Crithidia mellificae* Langrdige & Mcghee, a trypanosome, were found in high copy number in honey bees using HTS (Runckel et al., 2011). The virus and trypanosome were confirmed to be present in honey bees by other methods as well, but their relevance to honey bee health has yet to be determined.

Overall, the increasing power of molecular techniques allows us to exhaustively, accurately, and quickly ascertain what viruses and microorganisms infect honey bees. Information about the genome and transcriptome can also be obtained to identify genotypes and genes associated with the presence of these viruses and microorganisms. Using these molecular tools, I investigated a hypothesized immune response to picornaviruses, polyA rRNA, first reported by Johnson et al. (2009).

# 3. Materials and Methods

## 3.1 Comparison of Hawaiian virus-low bees and USDA virus-high bees using qRT-PCR

Thirty honey bees were obtained from Danbury Apiaries, a honey bee wholesaler in Kauai, Hawaii. Nurse bees were placed directly into liquid nitrogen and shipped to the University of Illinois at Urbana-Champaign and then stored at -80 °C until RNA extraction. The placement of bees directly into liquid nitrogen Dewar flasks limits any changes in gene expression during the collecting process. Thirty honey bees collected from the USDA laboratory in Beltsville, MD were placed in a dry ice container and then stored in -80 °C. In both cases, bees were preserved at extremely cold temperatures to maintain the integrity of honey bee RNA and viral RNA. All bees were collected in the summer of 2011.

In effort to simplify the process of polyA rRNA diagnosis, I avoided a gut dissection and instead used the entire abdomen, unlike Johnson et al. (2009), where total RNA was only extracted from the gut. Picornaviruses have been observed to be actively replicating in the whole abdomen in past studies (Fievet et al. 2006; Boncristiani et al. 2009). The inclusion of the entire abdomen however leads to the inclusion of plant material and other contaminants in the honey bee abdomen that interferes with extraction of pure RNA. Due to these contaminants, conventional RNA extraction methods that use RNeasy kits (Qiagen, Germany) and Trizol (Life Technologies, USA) could not be utilized in this project. Instead, an RNA extraction protocol designed for plant tissue RNA extraction developed by the Hudson lab was used to extract RNA from honey bee abdomens. Honey bee abdomens were separated from thoraxes over dry ice

and then placed immediately into 440 ul cetrimonium bromide (cetyl trimethylammonium bromide, CTAB) buffer heated to 65°C. CTAB buffer is designed to remove plant-related phenols that contaminate RNA extracts. Only six samples were done at a time to ensure that the CTAB buffer remained warm throughout the processing of all six samples. While in CTAB buffer, honey bee abdomens were ground with a motorized pestle for 30 s. To ensure thorough homogenization, I also used a homogenizer to grind the abdomens for an additional 30 s. The homogenizer tip was rinsed in double distilled water to flush out any honey bee remains between homogenizations. A phenol chloroform extraction was then performed. The aqueous layer was transferred to a new tube, to perform an ammonium acetate precipitation overnight. Afterwards, the mixture was spun at 14000 rpm for 30 minutes at 4°C. The pellet was washed with 80% ethanol and precipitated again with ammonium acetate for one hour. The pellet was washed again with 80% ethanol and then dissolved in 100 ul RNase-free water.

The quality of the RNA was checked with a Nanodrop spectrophotometer (Life Technologies, USA). RNA samples were purified for polyA sequences using Dynabeads Oligo (dT)25 magnetic beads (Life Technologies, USA). These magnetic beads use an attached oligo-dT to enrich for poly(A) sequences in a sample. The RNA quality of each sample was checked with an Agilent Bioanalyzer 2100 using an RNA 6000 Nano chip (Agilent Technologies, USA). Hymenopteran 28S RNA is known to undergo cleavage at low temperatures (Winnebeck et al., 2010). Thus, the Agilent RNA Integrity Number (RIN) was not used to measure the quality of RNA. Instead, the electropherogram produced by the Agilent Bioanalyzer was visually analyzed to check for any sign of RNA degradation (Dainat et al., 2011; Winnebeck et al., 2010). RNA was then reverse-transcribed into cDNA using random hexamers and Arrayscript reverse

transcriptase (Life Technologies, USA). Oligo-dT primers were tested as primers for cDNA synthesis, but random hexamers were chosen instead because the secondary structure of rRNA was considered less likely to interfere with cDNA synthesis if random hexamers were used.

Quantitative RT-PCR was performed on each bee with three technical replicates using FastStart Universal SYBR-Green master mix (Roche, Switzerland) on the Applied Biosystems Prism 7900HT System (Life Technologies, USA). Two reference genes, Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (GB14798) and  $\beta$ -actin (GB17681), were measured to standardized gene expression data (Scharlaken et al., 2008; Vandesompele et al., 2002). As a control for the reverse transcription reaction, 0.01 ng Arabidopsis root cap protein (AY056187) RNA was spiked into each sample. Primers from vanEngelsdorp et al. (2009) were used to measure DWV and BQCV quantities. Ribosomal primers were designed using the *A. mellifera* ribosomal sequences found in Gillespie et al. (2006). Primers developed by Jay Evans (USDA-ARS) and Thomas Newman (UIUC) were used in to test the ribosomal sequences noted in Johnson et al. (2009) (Table 1). PolyA rRNA expression level between virus-high and virus-low bees were compared using relative quantification ( $2^{-\Delta\Delta C_T}$  method) (Livak & Schmittgen, 2001). Absolute quantification was not applied in this study because of the high abundance of rRNA. Relative quantification is a simpler method to establish fold change gene expression between conditions, so absolute quantification is not necessary. The method to use real-time PCR data to calculate relative quantity is given below:

$$\Delta\Delta C_T = (C_{T,goi} - C_{T,ref})_{virus-high} - (C_{T,goi} - C_{T,ref})_{virus-low}$$

where *goi* represents the virus or ribosomal fragments and *ref* represents a reference gene. All samples were run in triplicate.

I compared *Varroa*-free and *Varroa*-infested bees to test for a significant difference in polyA rRNA and virus load using a Mann-Whitney U test to account for the expected high variation of rRNA and viral titers in honey bee RNA extractions. Additionally, I calculated the fold change in expression between the three groups with the following equation:

$$\text{Fold change} = 2^{-(C_{T,goi} - C_{T,ref})_{virus-high}} - 2^{-(C_{T,goi} - C_{T,ref})_{virus-low}} \text{ (Schmittgen \& Livak, 2008)}$$

### 3.2 Mining for polyA rRNA sequences in RNA-seq data

*A. mellifera* RNA-seq datasets were opportunistically obtained from past experiments completed by Khetani & Evans (unpublished data, 2012) and Alaux et al., (2009). The RNA-seq datasets were collected with two different methods, described below.

In Khetani & Evans (unpublished data, 2012), RNA-seq was used to measure changes in gene expression in response to *N. ceranae*. Total RNA was extracted from whole abdomens of individual worker bee samples and processed to be sequenced using the Illumina Genome Analyzer (Illumina, USA). Typical RNA-seq libraries include a polyA purification step using oligo-dT magnetic beads so that mRNA is enriched; however, this step was omitted in these library preparations so that the metagenome and transcriptome could both be analyzed. The study yielded 32 RNA-seq libraries, 16 of *Nosema*-negative bees and 16 of *Nosema*-positive bees. Khetani & Evans (unpublished data, 2012) also measured viral load for the bees in this study. Viral loads were used to separate bees into two groups: a virus-low group of bees with < 10,000 virus reads and a virus-high group of bees with > 10,000 virus reads per individual (Fig 1). The RNA-seq statistical analysis software DEGseq was also run to test if immunity genes were differentially expressed between these two groups (Wang et al. , 2010).

In Alaux et al (2009), the origin of the second set of RNA-seq data, RNA-seq was used to measure the difference in gene expression between *A. mellifera scutellata* (AHB) and *A. mellifera ligustica* (EHB). Total RNA was sequenced from brains of individual nurses and foragers using the Illumina Genome Analyzer. Both RNA-seq libraries generated 80 nt libraries. I used these RNA-seq data from the brain to act as a control to the RNA-seq data from the abdomen in Khetani & Evans (unpublished, 2012). I hypothesized that polyA rRNA is a response to viral infection, so I expected a lower polyA response in the brain compared to the abdomen because more viruses occur in the abdomen (Fievet et al. 2006; Boncristiani et al. 2009).

A third set of RNA-seq data was used to determine if similar polyA regions to *A. mellifera* could be found in *D. melanogaster*. I obtained *D. melanogaster* RNA-seq datasets in the NCBI Sequence Read Archive that were collected from a transcriptome study performed by Graveley et al. (2011). *D. melanogaster* whole body RNA was collected from adult females and males from three different life stages. RNA-seq libraries were created using the Illumina Genome Analyzer Ix platform, generating ~76 nt reads. If polyA rRNA is a response to picornavirus infection, then I would expect to see less polyA rRNA in the presumably healthy flies used by Graveley et al. (2011).

I searched for polyA rRNA reads in all RNA-seq datasets obtained from Khetani & Evans (unpublished, 2012), Alaux et al. (2009), and Graveley et al. (2011) using the same workflow. First, linker and adaptor sequences were removed using a Perl script created by the Hudson laboratory. Afterward, I used a Perl and shell script to identify polyadenylated reads from RNA-seq data and trim the polyadenylated tails from reads (Appendix A). Bowtie (Langmead et al., 2009) was used to align edited *A. mellifera* reads against an indexed database of the honey bee

ribosome (Gillespie et al., 2006) and *D. melanogaster* reads against an indexed database of the fruit fly ribosome (Tautz et al., 1988). Bowtie allows a maximum of two mismatches when mapping reads to the reference sequence, which is biologically sound in the case of comparing ribosomal sequences between intraspecific populations. Reads that were aligned to ribosomal sequences using Bowtie were then assembled to their respective rRNA sequence using the Geneious Assembler (Drummond et al., 2011).

After alignment to the rRNA sequence, polyA rRNA expression in regions with a high number of reads was analyzed between virus-low and virus-high groups in Khetani & Evans (unpublished data, 2012). PolyA rRNA read counts were normalized against total RNA to yield polyA rRNA reads per million total RNA reads (RPM) for individual bees. All regions found in the 18S rRNA were grouped to compare polyA rRNA RPM between virus-low and virus-high groups using an ANOVA and negative binomial model. The same analysis was done in the 28S rRNA to compare polyA rRNA RPM in regions with a high number of reads between virus-low and virus-high groups. When analyzing the 18S and 28S polyA regions individually, I used the Mann-Whitney U test to compare the amount of polyA rRNA RPM between virus-low and virus-high groups. RNA-seq datasets from Alaux et al. (2009) and Graveley et al. (2011) did not exhibit particular rRNA regions with a high number of reads, so the polyA rRNA RPM analysis was not performed for those datasets.

To account for possible sequence homology between *A. mellifera* polyA rRNA sequences and mRNA sequences, putative polyA rRNA sequences were searched for in Honey Bee Assembly version 4.5 on BeeBase with BLAST (Munoz-Torres et al., 2011). Ribosomal RNA

shares some homology with mRNA transcripts, but if most of the 80 nt read mapped to the rRNA, I assumed that the read is rRNA (Mauro & Edelman, 1997).

To measure possible template switching during reverse transcription, the ends of the aligned reads were analyzed for single nucleotide polymorphisms (SNPs) (Appendix B). Template switching is a phenomenon where a reverse transcriptase begins its activity on one transcript and then switches to another homologous template, creating cDNA different from the original RNA transcript (Cocquet et al., 2006; Gilboa et al., 1979). The transcripts found in this study could be chimeric reads from a product of template switching. To check for template switching, polyA ends of each were identified and then the bases near that end were measured for variability from the *A. mellifera* RNA sequence (Gillespie et al., 2006). If high variation was detected in the 3' ends of the reads, then the read could be chimeric an mRNA-rRNA read.

One could also expect chimeric viral RNA and *A. mellifera* rRNA reads if the reverse transcriptase hopped from virus RNA to *A. mellifera* RNA. To account for this possibility, BLAST (Zhan et al., 2000) was used to check if polyA rRNA was similar to ssRNA virus NCBI entries (taxid:439488), which includes the picornaviruses and nodaviruses (LSV).

### **3.3 PCR of polyA regions discovered in RNA-seq analysis**

The initial qRT-PCR assay was designed using several sets of primers that detect different regions of the ribosome based on Johnson et al. (2009). After finding polyA regions with a high number of reads with the opportunistically-gathered RNA-seq data, I attempted to design new primers to target that region using Primer3 (Rozen & Skalesky, 2000). Using this software, no primer sets could be created, most likely due to the short sequence of the polyA

regions. Instead of using software to design the primer sets, new primers were designed manually, making sure that the primer pairs had similar melting points and a low probability of self-annealing. I used these new primers to target polyA regions of the ribosome in a subset of the original 60 virus-low and virus-high bees from Maryland and Hawaii. Only the 28S rRNA 2070-2130 region was found to be significantly different and was tested again using the same qRT-PCR protocol mentioned above.

## 4. Results

Initial qRT-PCR results demonstrated that the virus-high bees from the USDA indeed had a higher viral titer compared to the virus-low bees from Hawaii (Fig 2 and 3). DWV and BQCV are both higher in putatively virus-high bees for both the GAPDH and  $\beta$ -actin reference genes ( $p < 0.001$ ). The 18S 291-359 region was higher in virus-high bees when using GAPDH as the reference gene ( $p < 0.05$ ) but was not significantly different when using  $\beta$ -actin as the reference gene ( $p = 0.771$ ) (Fig 4). GAPDH and  $\beta$ -actin produce different results when used to calculate relative expression, which may indicate that these reference genes may be manipulated by the conditions of the experiment. Expression data were log-transformed to account for the large differences in quantity of virus and ribosomes in samples.

The results of mining for polyA sequences in the three RNA-seq libraries (honey bee abdomen, honey bee brain, and fruit fly) show that polyA sequences made up less than 0.01% of the RNA-seq library in all cases (Table 2). PolyA 18S and 28S rRNA made up about 1% of polyA rRNA, which is much lower than the typical >50% rRNA that usually comprises total RNA. The 18S polyA rRNA showed higher coverage in three ~60 nt regions when mapped to reference rRNA sequence and the 28S polyA rRNA showed higher coverage in two ~60 nt regions (Table 3). These five regions were not clearly identifiable in honey bee brain and *D. melanogaster* RNA-seq libraries (Fig 5 and 6). For both 18S and 28S polyA rRNA, *D. melanogaster* had much shorter polyA rRNA reads than *A. mellifera*, further suggesting that polyA rRNA is a cellular response unique to the *A. mellifera* abdomen (Fig 7 and 8).

To confirm that polyA rRNA reads mapped best to rRNA sequences, I used BLAST to compare the nucleotide sequence of the unique 18S and 28S rRNA regions against Honey Bee Genome Assembly 4.5 and ssRNA virus entries in NCBI. No similarities were found for all regions when searched for on both databases. The fragments detected in Table 3 map best back to the ribosome, not other parts of the genome, which suggests that the polyA fragments detected in RNA-seq analysis are 18S and 28S rRNA.

It is a frequent observation in cDNA libraries that ribosomal sequences are present, and this can be explained by ligation chimerism and by reverse transcriptase template switching in addition to rRNA polyadenylation. I investigated the polyA rRNA reads for signatures of template switching or chimerism. The six bases following the polyA tail in all three 18S rRNA polyA regions are consistent with the rRNA reference sequence in Gillespie et al. (2006), suggesting that the reads in these regions are not chimeras but actually a fragment of 18S polyA rRNA sequences (Fig 11). However, two polyA regions in the 28S rRNA have two SNPs per region when compared to the rRNA reference sequence (Fig 12). These two SNPs could be taken to indicate chimeric reads, but the fact that only two nucleotide differences are observed is more likely to be accounted for by polymorphisms between that rRNA and the reference sequence or by sequencing or polymerase errors.

DEGseq analysis performed between the two groups revealed 52 differentially expressed genes between the two groups (Table 4). Specific immunity-related genes mentioned in Evans et al. (2006) were found to be downregulated in virus-high bees.

The amount of polyA rRNA in the specified regions in Table 3 was compared between virus-low and virus-high bees (Fig 9 and 10). No difference was found between the three

regions in the 18S rRNA and the 28S 35-93 region, but a marginally significant difference was found in the 28S 2070-2130 ( $p = 0.09$ ). In the 28S 2070-2130 region, virus-low bees have a median value of 2.39 polyA rRNA RPM while virus-high bees have a median value of 7.05 polyA rRNA RPM. The difference in expression levels of 28S 2070-2180 polyA rRNA between virus-low and virus-high bees was trending toward significance, so qRT-PCR for the region was performed on additional honey bees.

The 28S 2070-2130 region was investigated further by quantification of polyA rRNA in the previously examined virus-low bees from Hawaii and the virus-high bees from Maryland using qRT-PCR. In the initial qRT-PCR, the primers only tested for regions noted in Johnson et al. (2009). Primers used in the second round of PCRs were designed after the RNA-seq analysis performed in this study (Table 5). When comparing 28S 2070-2130 rRNA expression between virus-low and virus-high bees using GAPDH and  $\beta$ -actin as reference genes, virus-high bees exhibited slightly higher expression levels of polyA rRNA ( $p = 0.2475$  and  $p = 0.678$ , respectively) (Fig 13). As observed in the qRT-PCR of the 18S 291-359 region, the different results derived from the GAPDH and  $\beta$ -actin reference genes could be evidence for changes in expression of these genes in response to the conditions of the experiment.

# 5. Discussion

## 5.1 Initial qRT-PCR

Hawaiian bees collected as negative control bees had low titers of DWV. Inapparent infections of picornaviruses are common, and bees that that appeared phenotypically disease-free could still have harbored the virus (Carter & Genersch, 2008). Inapparent covert infections in both Hawaiian and USDA honey bees could confound control and experimental groups, making the comparison of viral titers between groups less certain. Aside from *V. destructor*, DWV can be transmitted between honey bees vertically and in larval food and horizontally through trophallaxis (Chen et al., 2006). Asymptomatic colonies, even those found in Hawaiian *Varroa*-free areas, may still be infected with picornaviruses because of these other routes for transmission of pathogenic viruses.

DWV and BQCV titer varied greatly among both sets of Hawaiian and USDA bees, a result that is similar to those of past attempts to quantify picornaviruses (Dainat et al., 2011). Honey bees were collected in the same season of the same year, so variation in DWV titer should be based wholly on the variation seen between honey bees of the same colony (Gauthier et al., 2007). Even within individual bees, DWV virions can highly variable between tissues (Boncristiani et al., 2009; Genersch et al., 2006; Yue & Genersch, 2005). Virion quantity can also depend upon infection stage of the host honey bee. Picornaviruses infect the alimentary canal of the honey bee and are subsequently shed through defecation (Miranda,

2008). Honey bees that are recovering from picornavirus infection would be categorized as highly infected when diagnosed with abdominal tissue due to high amounts of virus in the gut.

Other measures can be taken to quantify picornavirus titer more precisely in future experiments. Picornaviruses are positive strand RNA viruses that replicate via a negative strand intermediate, so designing primers for the negative sense strand of DWV and BQCV may have been a more precise way to analyze infection status (Boncristianiet al. 2009). Additionally, picornavirus titers in the *V. destructor* itself may be a better measure of DWV titer in honey bees because the mite does not pass the virus through its alimentary canal as quickly as honey bees (Gisder et al., 2009). DWV has been shown to actively replicate in mites. Measuring DWV titer in mites has been presented as a means to provide an estimate of the amount of actively replicating virus in the host (Ongus et al., 2004; Yue & Genersch, 2005).

## **5.2 RNA-seq analysis**

The presence of many polyA transcripts in specific regions of *A. mellifera* 18S and 28S rRNA supports the discovery of polyA rRNA in Johnson et al. (2009). The abdomen has signs of polyA rRNA in at least five different regions. The RNA-seq analysis also clarified the sequence of the polyA rRNA, which was not done in Johnson et al. (2009). Although the probe sequence on the microarray was known in Johnson et al. (2009), the sequence of the putative polyA rRNA transcript was not confirmed with sequencing. In this project, the RNA-seq analysis identified the polyA rRNA sequence. However, the regions found in this analysis differ from those found in Johnson et al. (2009). In the initial qRT-PCR, the 18S 291-359 rRNA region that exhibited

higher expression in virus-high bees could have been only measuring the noise in regions of the polyA 18S rRNA (Fig 4 and 5).

The low number of polyA reads can be explained by the type of extraction and reverse transcription. RNA was enriched for polyA sequences using oligo-dT beads, but ribosomal RNA was not completely subtracted from the sample. When the total RNA was sequenced, the sample had a large percentage of rRNA, thereby reducing the amount of polyA RNA sequenced. In addition, random hexamers, not oligo-dT primers, were used to prime reverse transcription before sequencing. Random priming will not always target the polyA tail, so the 80 nt reads created by the Illumina Genome Analyzer will begin at random places in RNA transcripts. While polyA sequences are common in mRNA, these two factors combine to explain the low percentage of polyA sequences found in the RNA-seq libraries.

The scattered and random placement of polyA reads in the total RNA of the brain suggest that no specific, consistent polyadenylation sites are predominantly utilized in the rRNA of the brain as they are in the abdomen. The virology of ssRNA viruses also agrees with the mapping data in Figure 5 and 6. All known *A. mellifera* viruses are found in the abdomen of the honey bee, with the exception of Kakugo virus, which is found in the brain (Fujiyuki et al. 2006; Chen & Siede 2007). Ribosomal RNA in the brain may not be responding to infection because few or no viruses are present. Cells in other tissues in the honey bee, such as the abdomen, are likely to be exposed to more pathogens. Another possibility is that the viruses are present in the brain, but the brain lacks the means to respond effectively to the infection via ribosomal polyadenylation.

The *D. melanogaster* samples investigated do not exhibit the specific rRNA polyA sequences seen in *A. mellifera* (Figure 2). PolyA rRNA reads in *A. mellifera* are longer, more specific, and more abundant compared to the polyA rRNA reads found in *D. melanogaster*. Only two non-lethal ssRNA viruses have been found in *D. melanogaster*, Nora virus and *Drosophila C* virus, whereas eighteen lethal and virulent ssRNA viruses have been found in *A. mellifera* (Habayeb et al. 2006; Chen & Siede 2007; Hedges & Johnson 2008). PolyA rRNA may be found in honey bees due to the increased likelihood of ssRNA virus infection in comparison to *D. melanogaster*, or alternatively *A. mellifera* may possess a defense mechanism that *D. melanogaster* lacks.

SNPs were not detected in the 18S rRNA but some variation was seen in the 28S rRNA. This finding suggests that reverse transcriptase template switching from mRNA or virus RNA transcript to rRNA transcript did not occur in the 18S rRNA regions. The variation at the ends of the 28S rRNA polyA reads may cast some doubt that they are true rRNA reads. During cDNA synthesis, the reverse transcriptase may have started on a polyA transcript, for example an mRNA transcript, and switched templates to rRNA. When this possible chimera was sequenced, it would have been falsely identified as a polyA transcript. Ribosomal RNA is so abundant in the cell that switching can occur and could be the reason for the presence of rRNA on previous microarrays. If template switching in this region is actually occurring, then picornaviruses could play a role in template switching between mRNA transcripts and 28S rRNA 2070-2130 fragments. However, the switching must be occurring with a transcript only two nucleotides different in sequence, and thus it is unlikely that such a sequence is not also a 28S rRNA. It is not certain why the 28S rRNA 2070-2130 is only fragment affected by picornaviruses, but this

novel response may only target one region in the ribosome. Only two out of the six last bases differed, so while there is a small possibility that the reads are chimeric, it is likely that they are actually rRNA reads.

No strong significance differences were found in 18S polyA rRNA RPM between the virus-low and virus-high groups. The cutoff at 10,000 reads, while initially arbitrary, was validated by DEGseq results. The downregulation of immunity-related genes in bees with greater pathogen load is similar to results from past research (Yang & Cox-Foster 2005; Chaimanee et al. 2012) and confirms that the *Varroa* load and/or the level of viral infection in the virus-high bees was sufficient to induce immunosuppression and distinguish them from virus-low bees.

Both the RNA-seq analysis and qRT-PCR have slight evidence that the 28S rRNA 2070-2130 region may be polyadenylated in response to viral infection. The association between picornavirus infection and this region of 28S polyA rRNA is trending toward significance more in the RNA-seq analysis rather than in the qRT-PCR analysis. In most cases qRT-PCR better quantifies genes that RNA-seq, but because the PCR product was not sequenced in this project, the RNA-seq analysis may offer better insight into the quantity of polyA rRNA between the virus-low and virus-high bees because the identity of the sequence is also known. Still, the RNA-seq analysis only offers some tenuous support of the association between polyA rRNA and viral load. While the identity of these sequences is known, the association of these polyA areas with pathogens or other physiological changes may be difficult to ascertain. Perhaps analysis with multiple viruses and polyA rRNA may give greater insight in the association between

picornaviruses and this immune response. Infection by multiple viruses, as suggested in Johnson et al. (2009), may play a stronger role in expression of polyA rRNA.

### 5.3 Conclusion

This study gives strong evidence for the presence of polyA rRNA fragments in virus-infected bees, as first suggested by Johnson et al. (2009). While the association between viral load and polyA rRNA expression level was not firmly established, the use of polyadenylation as a signal for the cell to degrade the ribosome may play a role in insect immunity. Cellular defenses to RNA virus infection by inhibition of translation have been observed in the past in HeLa cells (Bonderoff et al., 2008). Viral IRES elements undergo strong selection to have mutations that change its secondary structure, which could imply that the viruses are adapting to ribosomes (Martínez-Salas et al. 2001; Fernández-Miragall & Martínez-Salas 2007). Polyadenylation of rRNA could add to the suite of cellular defenses that inhibit viral replication in the cell.

Nonetheless, no strong association was found between virus load and polyA rRNA in this study. These rRNA markers were first found in association with CCD, but they can be associated with a number of debilitating conditions that other pathogenic and environmental stresses (Johnson et al., 2009). In future research, honey bees with other pathogens associated with CCD, such as IAPV and *N. ceranae*, can be tested for upregulation of the observed polyA rRNA fragments. As noted earlier, these markers arise from degradation intermediates from intracellular processes that stress initiates. Honey bees can endure stress from any number of factors, including parasites, insecticides and poor cultural practices (Oldroyd, 2007). PolyA rRNA

fragments were initially suggested as a marker for CCD, but they could also be general stress markers. In either case, the rRNA markers can be used to measure stress, CCD or otherwise, in which case they can be used when beekeepers are concerned about the health status of their bees.

To minimize the effect of other pathogens and environmental conditions, the RNA-seq analysis should be performed again with honey bees from *Varroa*-free and *Varroa*-infested islands of Hawaii. In the RNA-seq dataset used to compare polyA rRNA between virus-low and virus-high honey bees, half of the bees were inoculated with *N. ceranae*, which could have affected the outcome of the attempted association between polyA rRNA and virus load. DWV abundance is negatively associated with *N. ceranae* spore loads (Costa et al., 2011). To this date, BQCV abundance has a positive association with *N. apis* infection, but not *N. ceranae* infection (Bailey et al., 1983). Whether or not honey bee virus and *N. ceranae* quantities have a positive association, *N. ceranae* still presents a confounding factor in this experiment. Bees infected with *N. ceranae* have an altered immune response, so the polyA response expected in this study may have been suppressed (Antúnez et al. 2009; Chaimanee, et al. 2012). Future experiments testing for polyA rRNA should ideally use honey bees that are infected only with picornaviruses and no other malady in order to best control for immunosuppression.

Honey bee cell lines may be a better test of the significance polyA rRNA in disease progression. Cell culture offers a more controlled environment where viral quantity and cell quantity can be better measured. Protocols for creating and maintaining honey bee cell lines exist. Isolated viruses, on the other hand, do not exist and are required for proper rigorous testing of viral infection according to Koch's postulates (Hunter, 2010).

Genomic studies can help to identify genes for selective breeding that will confer a better immune response to viral infection. Gene expression studies reveal which genes play a role in pathogen defenses, such as the hypothesized polyA rRNA response. After the genes that confer immune responses are identified, beekeepers can focus their breeding efforts on bees that react to viruses. Currently beekeepers use acaricides to control *Varroa* and other chemicals to control for other threats to honey bee health. In the long term, selecting for bees with stronger immune defenses will ensure the future of honey bee health (Moritz & Evans, 2008).

## 6. Tables and Figures

Table 1. Primer sequences used for initial qRT-PCR. The 18S 291-359 primer set was originally used in Johnson et al. (2009) to measure the amount of polyA rRNA between honey bees from colonies diagnosed with CCD and healthy bees.

Sequence Description	Sequence (forward in first row, reverse in second row)	Bases	Length of amplicon	Source
18S 291-359	GTGCTGATCGCATGGTCATC	20	70	Thomas Newman, unpublished
	CCATGAACAGTTGATAAGGCAGAA	24		
Deformed Wing Virus (NC004830)	ATCAGCGCTTAGTGGAGGAA	20	130	vanEngelsdrop et al. 2009
	TCGACAATTTTCGGACATCA	20		
Black Queen Cell Virus (NC_003784.1)	GCAAGCTCTTCCAATGATAG	20	140	vanEngelsrop et al. 2009
	AAGATTCAGCCGAGTCCTT	19		

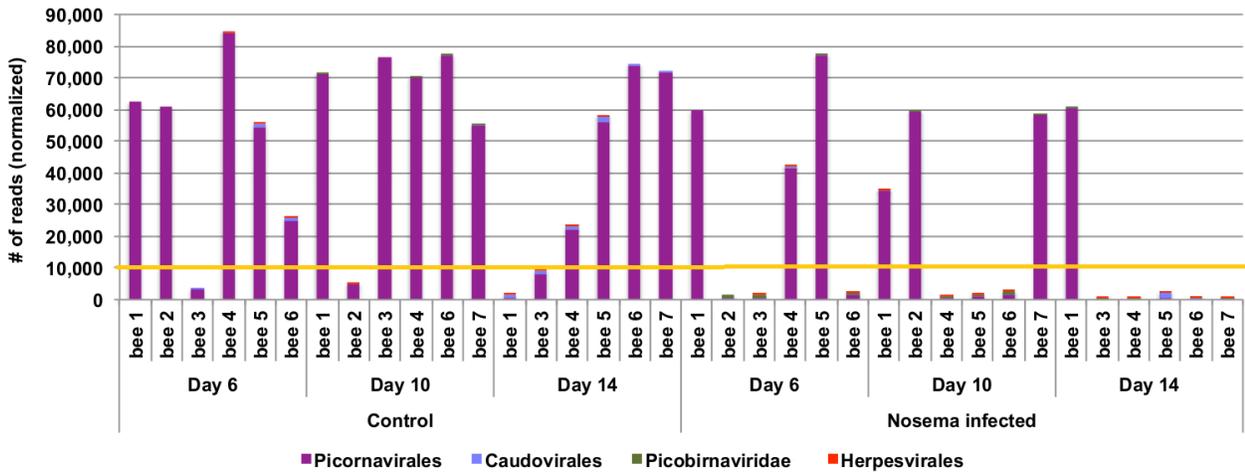


Figure 1. Viral counts in bees in a past experiment measuring changes in gene expression in response to *N. ceranae* infection. The division between virus-low and virus-high groups was made at 10,000 viral reads (yellow line), giving 14 virus-low and 22 virus-high honey bees. Figure generated by Khetani & Evans (unpublished data, 2012).

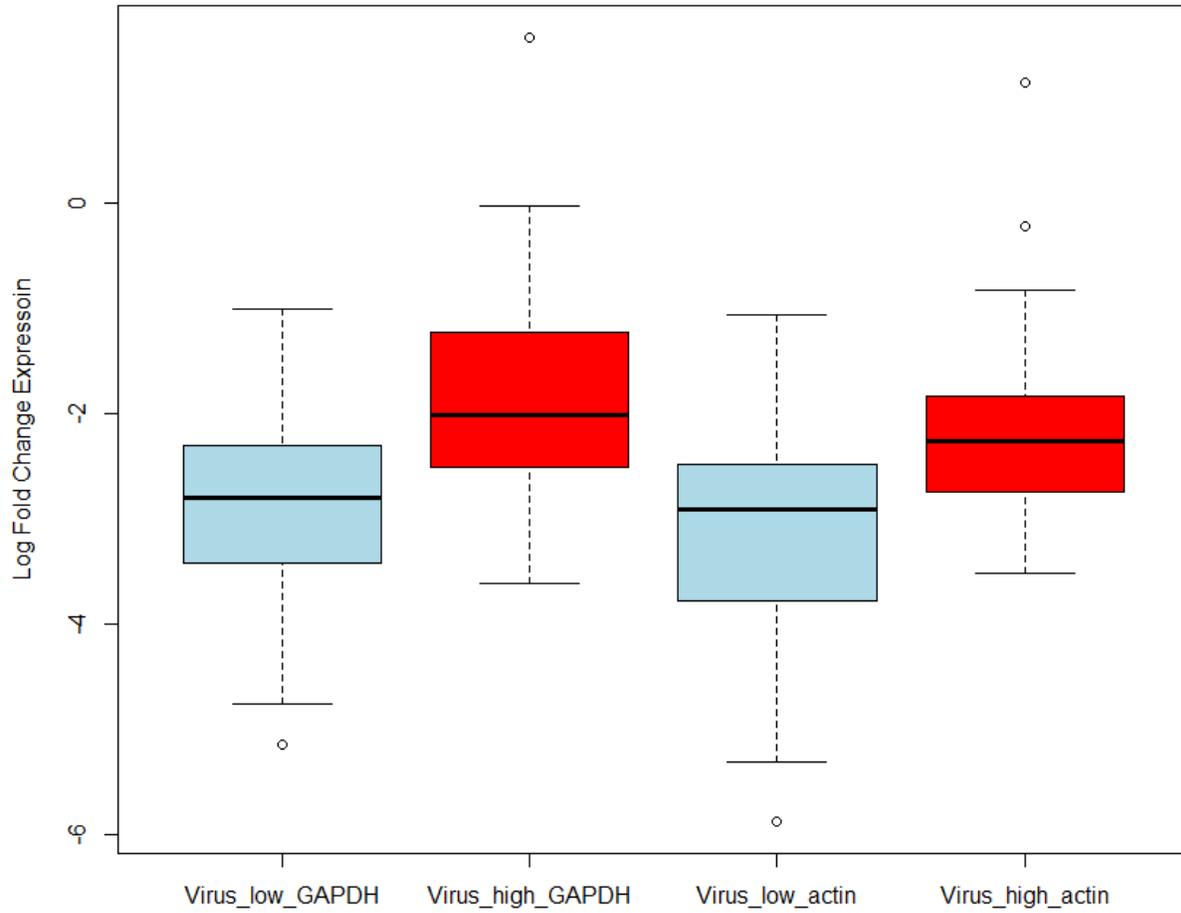


Figure 2. Log-transformed relative expression data of BQCV normalized against two reference genes, GAPDH and  $\beta$ -actin, between virus-low and virus-high bees ( $p < 0.001$  for both reference genes). The comparison of virus titer between bees from Hawaii and Maryland confirm that the Hawaiian bees have a lower virus titer as designed.

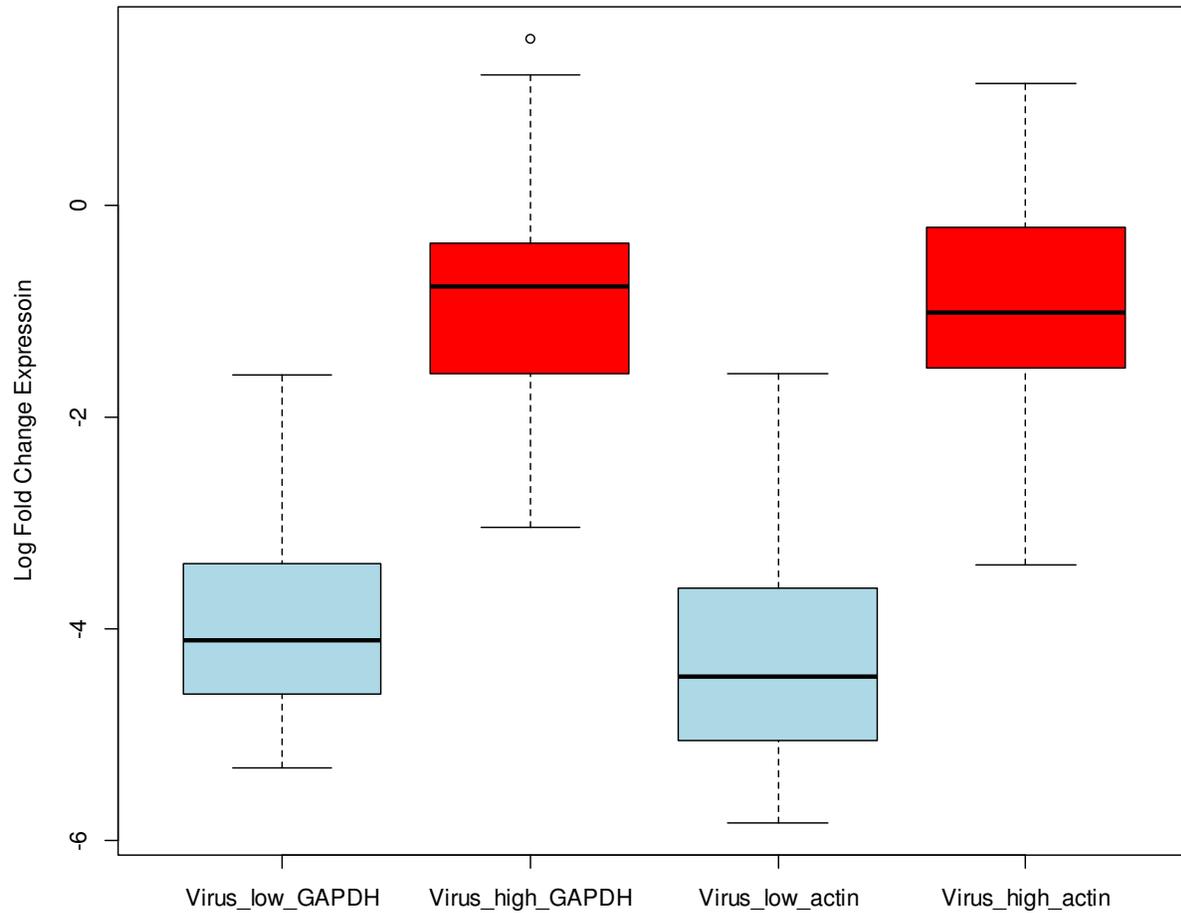


Figure 3. Log-transformed relative expression data of DWV normalized against two reference genes, GAPDH and  $\beta$ -actin (actin), between virus-low and virus-high bees ( $p < 0.001$  for both reference genes). As with the previous analysis of BQCV, the comparison of DWV titer between bees from Hawaii and Maryland confirms that the Hawaiian bees indeed have a lower DWV titer. The significantly different virus loads in the two groups of bees makes them suitable for comparison of the possible polyA rRNA immune response.

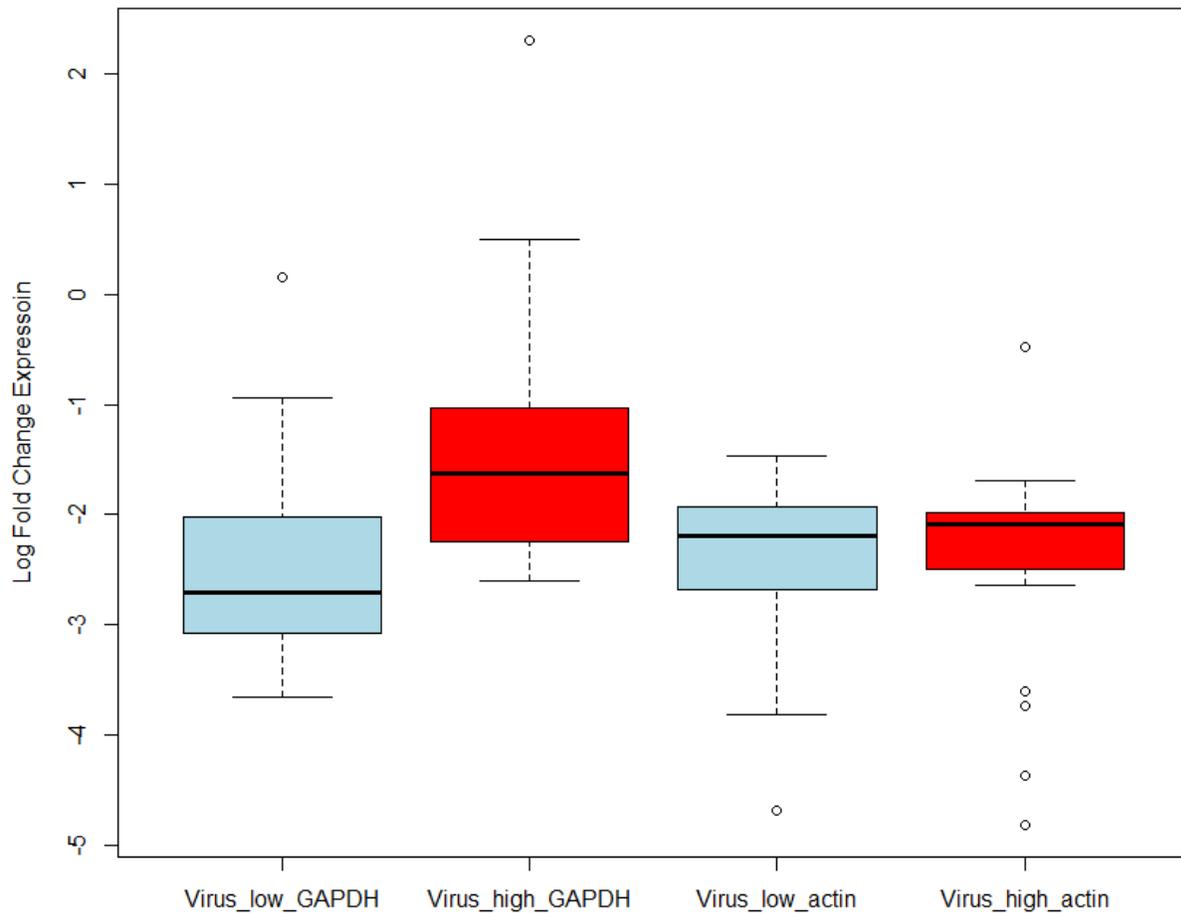


Figure 4. Log-transformed relative expression data of viruses and 18S 291-359 rRNA normalized against two reference genes, GAPDH and  $\beta$ -actin (actin), between virus-low and virus-high bees ( $p < 0.05$  for GAPDH,  $p = 0.771$  for  $\beta$ -actin). The RNA-seq analysis later showed this region did not undergo a high degree of polyadenylation. Even though polyA 18S rRNA was more highly expressed in virus-high bees when using GAPDH as the reference gene, the weak difference when using  $\beta$ -actin as the reference gene in combination with the low amount of polyA rRNA in this region suggest that this region of 18S rRNA does not participate in the hypothesized immune response.

Table 2. Percent polyadenylation in RNA-seq libraries in *A. mellifera* from a past *Nosema* inoculation experiment (Khetani & Evans, unpublished data, 2012), AHB/EHB gene expression experiment (Alaux et al., 2009) and transcriptome sequencing of *D. melanogaster* (Graveley et al., 2011). %PolyA shows the percentage of reads that are polyadenylated in the RNA-seq library. PolyA reads make up a small percentage of the library, as can be expected with libraries created with random hexamers primers. %18S and %28S represent the amount of polyA 18S and 28S found in total polyA sequences.

Sample	n	tissue	% PolyA	% 18S	% 28S
<i>Nosema</i>	36	whole abdomen	0.0194	1.04	1.10
AHB/EHB	32	brain	0.0155	0.495	0.543
<i>Drosophila</i>	22	whole body	0.0511	0.242	0.685

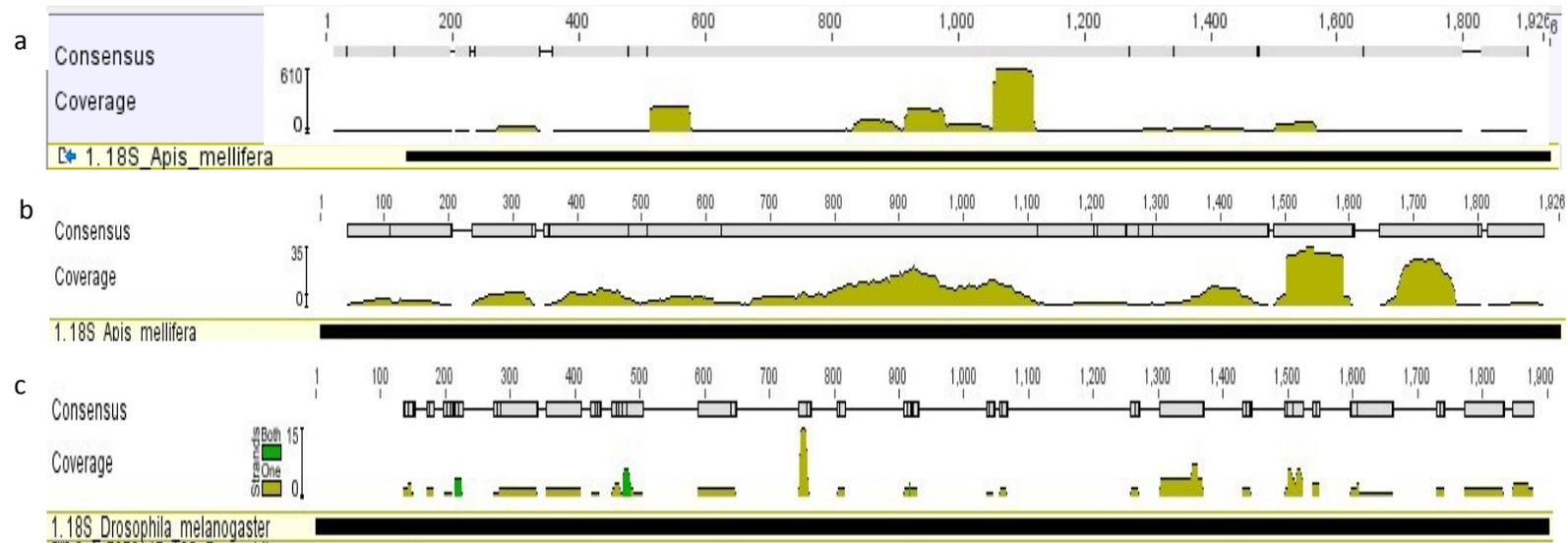


Figure 5. Location of polyA reads in the 18S ribosome. Width of bars represents length of sequence. In *A. mellifera* abdominal tissue (a), significant coverage was found in three regions in the 18S rRNA (514-576, 922-976, 1056-1120). PolyA reads in the abdomen map to a narrower range than polyA reads in the brain (b), suggesting that the abdominal cells have conserved polyA rRNA sites whereas the brain does not. Similarly to the honey bee brain, the scattered polyA regions in *D. melanogaster* (c) suggest a lack of conserved polyA sites.

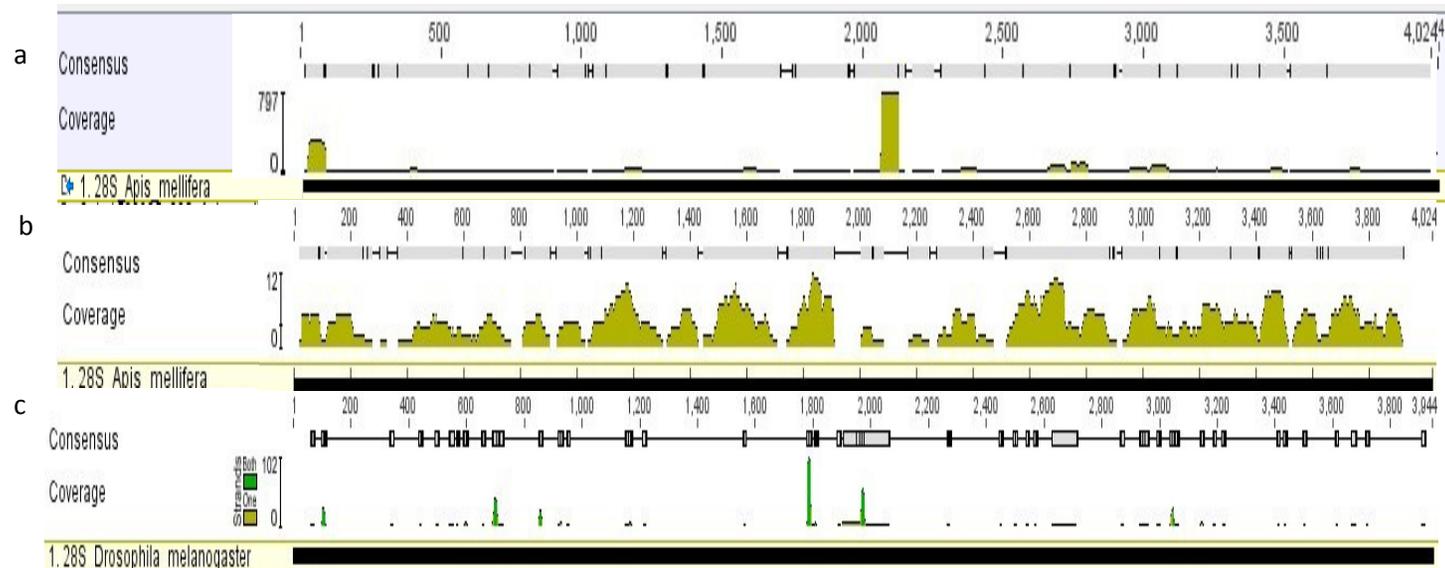


Figure 6. Location of polyA reads in the 28S ribosome. Width of bars represents length of sequence. In *A. mellifera* abdominal tissue (a), significant coverage was found in two regions in the 28S rRNA (35-93 and 2070-2130). Contrary to the *A. mellifera* abdomen, polyA rRNA reads were low in number and did not group to specific regions in *A. mellifera* brain (b) and *D. melanogaster* whole bodies (c). Like the 18S rRNA, only 28S rRNA in the *A. mellifera* abdomen appears to have conserved polyA sites, perhaps because of the greater incidence of picornavirus infection in that tissue.

Table 3. Sequences of polyA regions in *A. mellifera* ribosome derived from Figures 5 and 6.

Description	Sequence (5' to 3')	Length
18S 514-576	CGATACGGGACTCATCCGAGGCCCGTAATCG GAATGAGTACACTTTAAATCCTTTAACGAGG	63
18S 922-976	CGGAACCCCGAGGTAATGATTAATAGGGACAG ATGGGGGCATTCGTATTGCGACGTTA	58
18S 1056-1120	GTTAGAGGTTTCGAAGGCGATCAGATACCGCCC TAGTTCTAACCATAAACGATGCCAGCCA	60
28S 35-93	TTTAAGCATATTAATAAGCGGAGGAAAAGAAA CTAACAAGGATTTCTTAGTAGCGGCG	59
28S 2070-2130	AGAGTGTTTCGTCGGGGTAACCCAAAATGACCT GGAGACGCCGTCGGGAGATCCGGGAAGGG	61

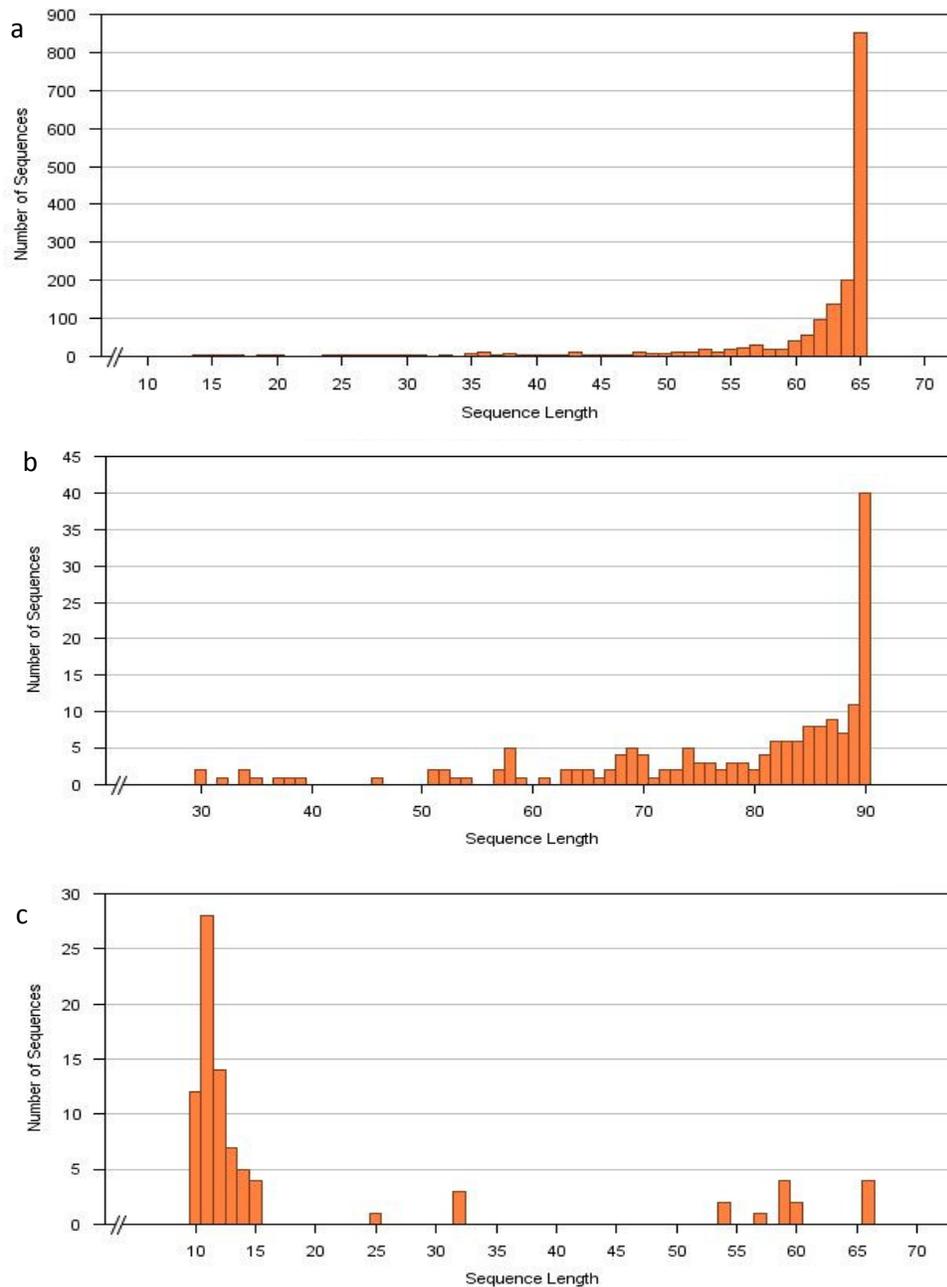


Figure 7. Read length of 18S polyA rRNA in *A. mellifera* and *D. melanogaster* in RNA-Seq datasets after removal of the polyA tail. The amount of polyA 18S rRNA reads ( $n = 1706$ , mean = 60.8) in the *A. mellifera* abdomen (a) is greater than the amount found in *A. mellifera* brains (b) ( $n = 178$ , mean = 77.6). The read length is greater in *A. mellifera* abdomen than in *D. melanogaster* whole body (c) RNA-seq datasets ( $n = 87$ , mean 19.8) ( $p < 0.005$ ). The low number of reads in *A. mellifera* brain and the short reads in *D. melanogaster* along with their scattered dispersal across the ribosomal RNA suggest that the abdomen has higher levels of 18S polyA rRNA.

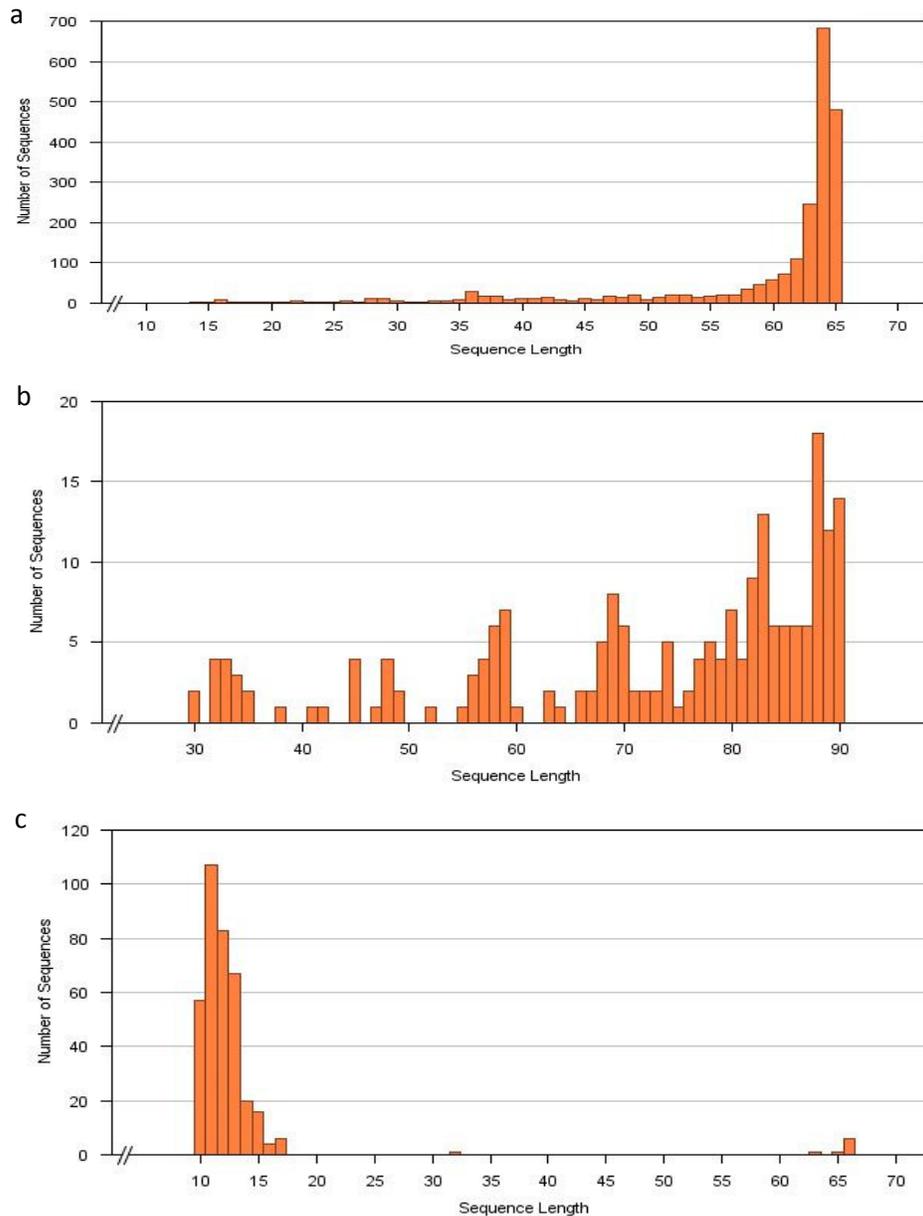


Figure 8. Read length of 28S polyA rRNA in *A. mellifera* and *D. melanogaster* RNA-Seq datasets after removal of the polyA tail. As with 18S polyA rRNA, the *A. mellifera* abdomen (a) ( $n = 1706$ , mean = 60.8) has more 28S polyA rRNA reads than the brain (b) ( $n = 206$ , mean 72.5). Additionally, the *A. mellifera* abdomen has 28S polyA rRNA reads that are greater in length than the reads found in *D. melanogaster* (c) ( $n = 369$ , mean = 13.2) ( $p < 0.005$ ). *A. mellifera* abdominal cells have greater read length and higher number of 28S polyA rRNA reads compared to brain cells and *D. melanogaster* cells, which suggests that detection of polyA rRNA is best done with the abdomen.

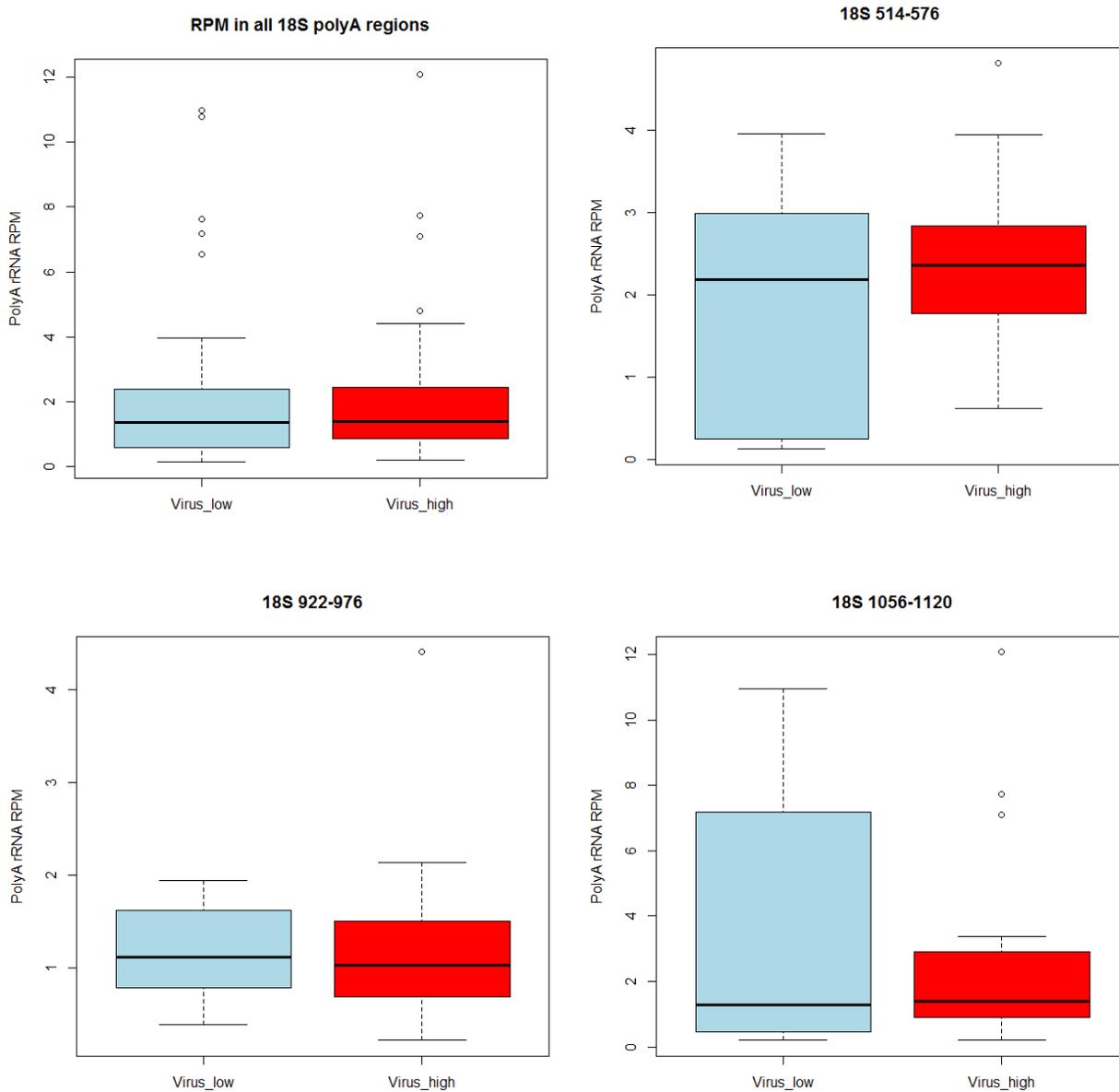


Figure 9. Boxplot of *A. mellifera* 18S rRNA polyA rRNA RPM. When grouped, the three polyA regions in 18S rRNA have similar polyA RNA RPM between virus low and virus high groups ( $p = 0.54$  ANOVA;  $p = 0.591$  negative binomial). RPM is not significantly different between virus-low and virus-high groups when regions are compared individually ( $p = 0.35$ ,  $p = 0.84$ , and  $p = 0.37$  in regions 514-576, 922-976, and 1056-1120, respectively). No significant association between 18S polyA rRNA and viral infection can be established.

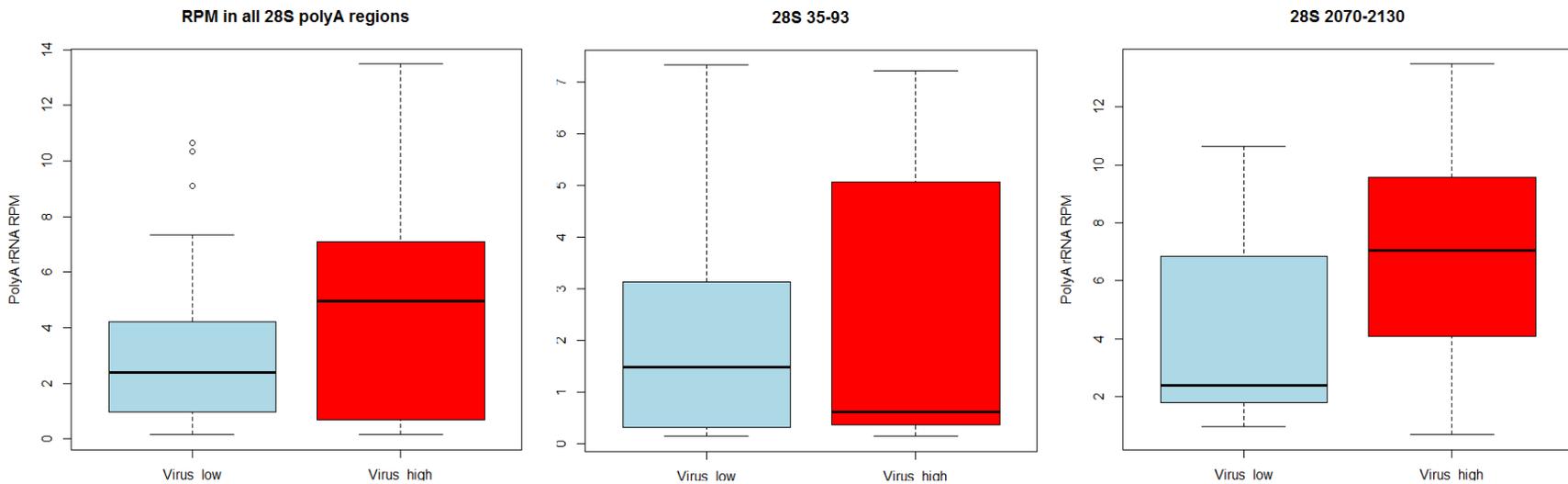


Figure 10. Boxplot of *A. mellifera* 28S rRNA polyA rRNA RPM. When grouped, the two polyA RNA regions are not expressed at significantly different levels between virus low and virus high groups ( $p = 0.155$  ANOVA;  $p = 0.176$  negative binomial). When comparing regions individually, the 28S 35-93 region exhibited the opposite trend, but still no significant difference was detected ( $p = 0.35$ ). The difference in polyA rRNA RPM between virus-low and virus -high groups is trending toward significance in the 28S 2070-2130 region ( $p = 0.09$ ). The 28S 2070-2130 region was tested more thoroughly with qRT-PCR to clarify the possible difference in polyA rRNA between virus-low and virus -high bees

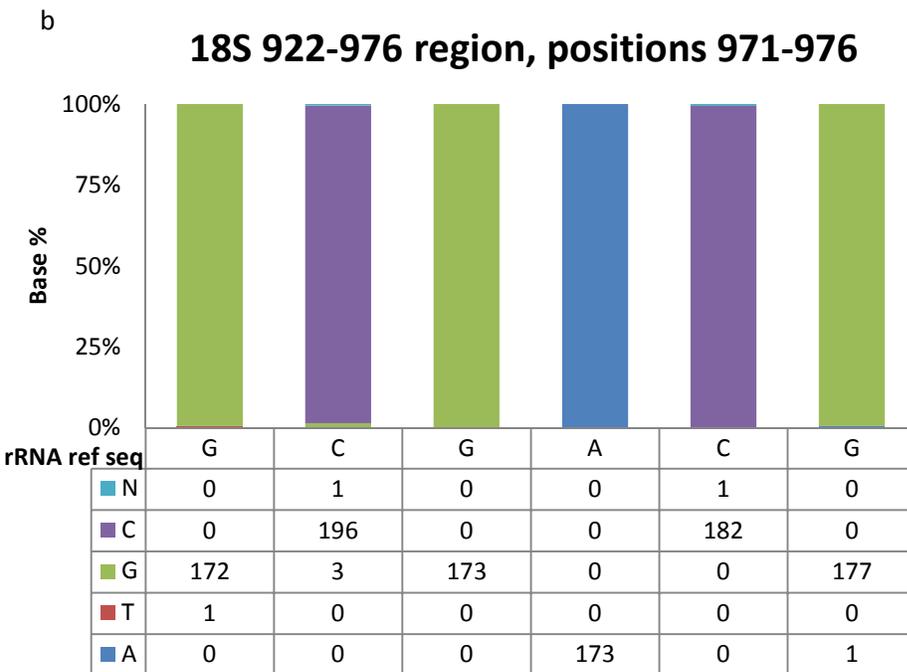
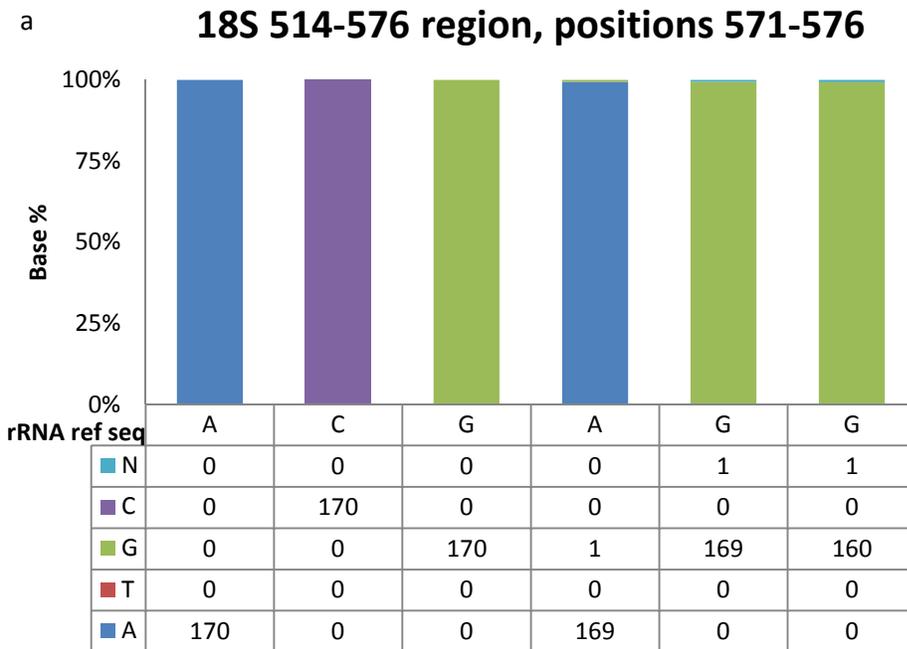


Figure 11. Nucleotide variation in the last six positions of the 18S polyA regions. 18S polyA regions (a-c) are homologous to the reference rRNA. The consistency between the reads and the reference rRNA indicates that template switching from an mRNA to rRNA most likely did not occur.

**c 18S 1056-1120 region, positions 1115-1120**

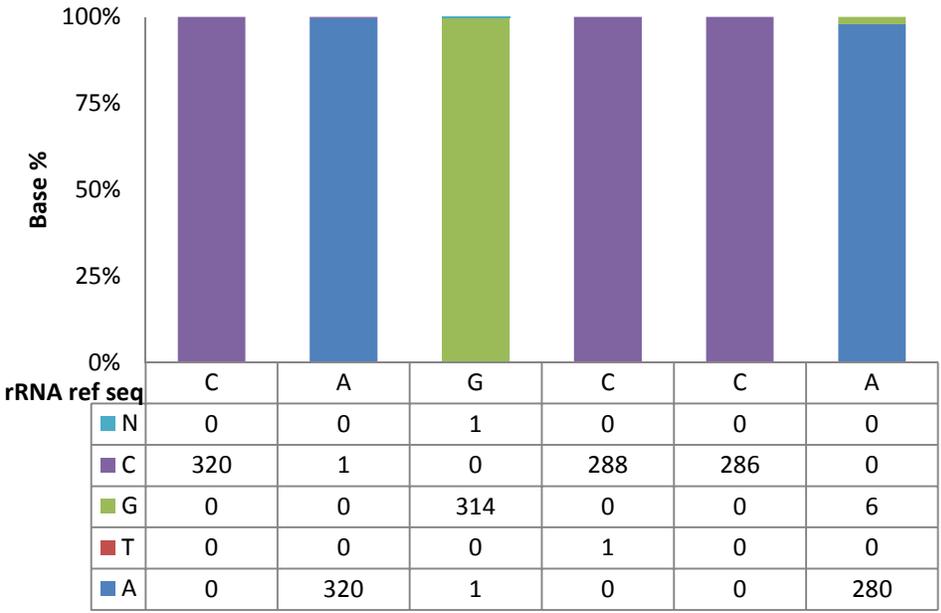


Figure 11 (cont.)

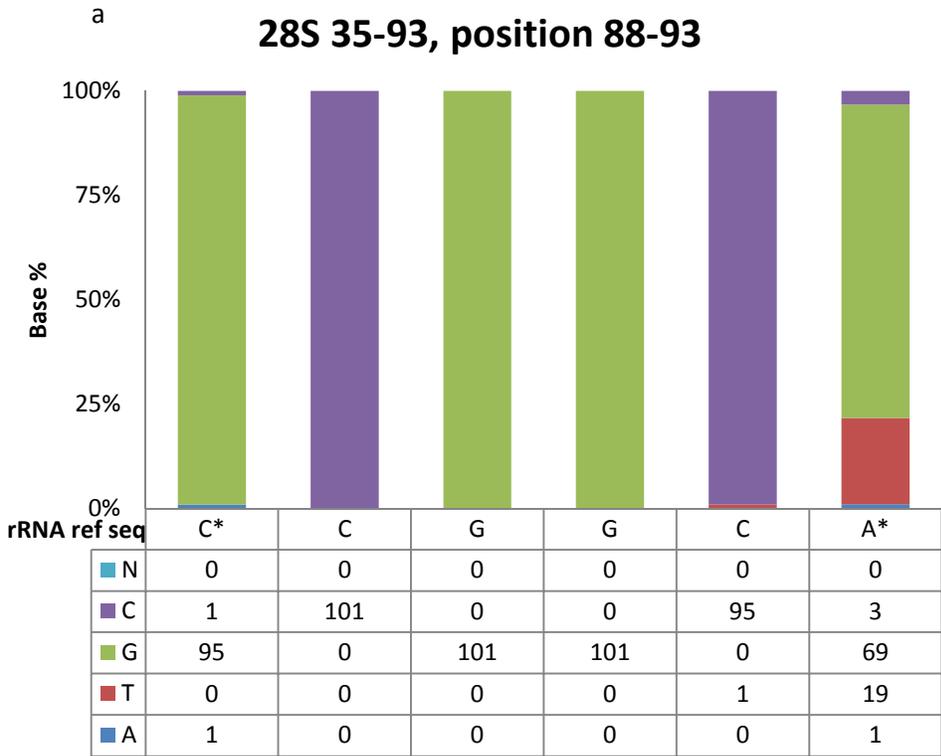


Figure 12. Nucleotide variation in the last six positions of the 28S polyA regions. Both 28S polyA regions (a-b) show variation in two positions in the last six bases. Positions with the majority nucleotides deviating from the reference rRNA sequence are denoted with an asterisk. The variability observed in the last six bases of 28S polyA rRNA could potentially indicate that the reads are not entirely rRNA but instead are chimeras of mRNA and rRNA that resulted from template switching during reverse transcription. However, the majority of the 28S region matches the honey bee 28S ribosome sequence in Gillespie et al. (2006), suggesting that these polymorphisms were only errors in sequencing or actual SNPs between honey populations.

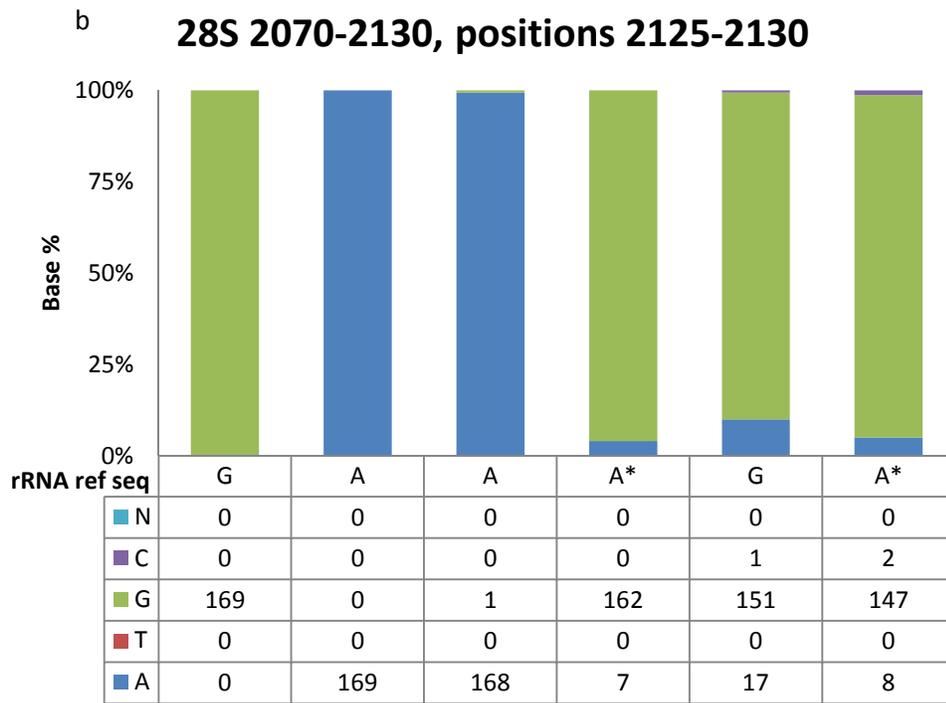


Figure 12 (cont.)

Table 4. Differentially expressed immunity-related genes found in DEGseq analysis. A total of 52 genes were found to be differentially expressed between the virus-low and virus-high bees. All antimicrobial proteins were found to be downregulated in virus-high bees, a response associated with acute viral infection, which provides validation for the distinction drawn between the virus-low and virus-high groups in the data provided by Khetani & Evans (unpublished, 2012).

Description	GB accession no.	log2FoldChange	pval
<i>Immunity</i>			
abaecin precursor	GB47318-RA	-0.892	8.90E-12
apidaecin	GB47546-RA	-0.362	9.39E-05
apidaecin precursor	GB51223-RA	-1.346	6.51E-13
irp30	GB50423-RA	-0.743	7.95E-08
protein g12	GB43688-RA	-1.849	1.28E-05
sestrin-1-like isoform 1	GB49564-RA	1.415	3.22E-05
lysozyme 3-like	GB47104-RA	1.009	1.26E-04
arsenite-resistance protein	GB43113-RA	1.159	1.46E-05
<i>Digestive</i>			
chitinase 5 precursor	GB53565-RA	-1.441	1.49E-06
chymotrypsin-2	GB45565-RA	-0.887	4.41E-05
<i>Metabolism</i>			
cytochrome o ubiquinol oxidase subunit iii	GB55956-RA	2.040	2.28E-07
cytochrome oxidase subunit i	GB53212-RA	0.776	8.53E-05
lipid storage droplets surface-binding protein 1	GB47140-RA	0.715	1.61E-04
lipoprotein	GB53878-RA	2.413	1.22E-05
arginine biosynthesis bifunctional protein	GB47897-RA	5.336	1.27E-05
invertase recombinase-like protein	GB43994-RA	4.357	6.86E-07
sodium-dependent phosphate transporter	GB47055-RA	1.101	7.65E-06
major royal jelly protein 1	GB55729-RA	1.309	3.03E-05
<i>DNA maintenance</i>			
dna helicase c2	GB55954-RA	2.555	1.55E-04
dna-directed rna polymerase subunit omega	GB54616-RA	3.145	2.77E-08
ribonuclease d	GB55126-RA	3.201	3.35E-07
ribonuclease e	GB46830-RA	2.586	1.23E-11
endoribonuclease dicer-like	GB48923-RA	0.964	7.39E-05

Table 4 (cont).

Description	GB accession no.	log2FoldChange	pval
<i>Possible xenobiotic response</i>			
aldo keto reductase	GB40571-RA	1.610	2.57E-08
probable cytochrome p450 6a14-like	GB40286-RA	1.870	1.63E-05
alkyl hydroperoxide reductase c22 protein	GB42133-RA	-0.760	6.77E-05
superoxide dismutase	GB45989-RA	2.779	3.38E-14
flavin-containing monooxygenase fmo gs-ox-like 3-like	GB45609-RA	0.817	3.57E-05
<i>Structural</i>			
endocuticle structural glycoprotein bd-1	GB46297-RA	0.636	1.67E-04
<i>Cellular communication</i>			
inositol-1-monophosphatase	GB50311-RA	4.486	3.69E-10
venom serine carboxypeptidase-like	GB40141-RA	-0.587	6.74E-05
<i>Miscellaneous genes</i>			
ets translocation variant 1	GB50156-RA	0.774	1.60E-04
hypothetical protein APA01_17930 [Acetobacter pasteurianus IFO 3283-01]	GB47865-RA	2.492	5.22E-05
hypothetical protein CIN_03810 [Commensalibacter intestini A911]	GB55863-RA	2.755	1.23E-04
low quality protein: frizzled-like	GB47061-RA	1.030	1.79E-04
major facilitator superfamily domain-containing protein 6-like	GB50150-RA	1.751	1.20E-07
outer membrane protein	GB43995-RA	1.220	6.70E-05
phospholipase membrane-associated-like	GB42475-RA	1.483	1.48E-05
poor imd response upon knock-in	GB44455-RA	1.660	6.79E-06
protein yellow-like	GB51759-RA	1.350	1.11E-04
signal peptide	GB47757-RA	3.025	1.29E-13
trigger factor	GB42389-RA	3.816	1.86E-06
uncharacterized protein loc100865648	GB42173-RA	2.754	7.85E-05
uncharacterized protein loc100865787	GB46427-RA	-3.335	6.72E-06
uncharacterized protein loc100877653	GB49655-RA	2.232	1.60E-04
uncharacterized protein loc100883608	GB50550-RA	1.949	6.01E-08

Table 5. Primers developed for qRT-PCR after RNA-seq analysis. Only the 28S 2070-2130 region showed a possible association between viral infection and polyA rRNA, so only that region was tested in the subsequent qRT-PCR.

Description	Forward	Reverse	Product Size (bp)
18S 514-576	CGATACGGGACTCATCCGA	CCTCGTTAAAGGATTTAA AGTGACTCA	63
18S 922-976	CGGAACCCCGAGGTAATGAT	TAACGTCGCAATACGAAT GCC	58
18S 1056-1120	GTTAGAGGTTCGAAGGCGATC A	TGGCTGGCATCGTTTATG G	60
28S 35-93	AGCATATTAATAAGCGGAGGA AAAG	CGCCGCTACTAAGGAAAT CCT	55
28S 2070-2130	AGAGTGTTTCGTCGGGGTAACC	GACGGCGTCTCCAGGTCA T	44

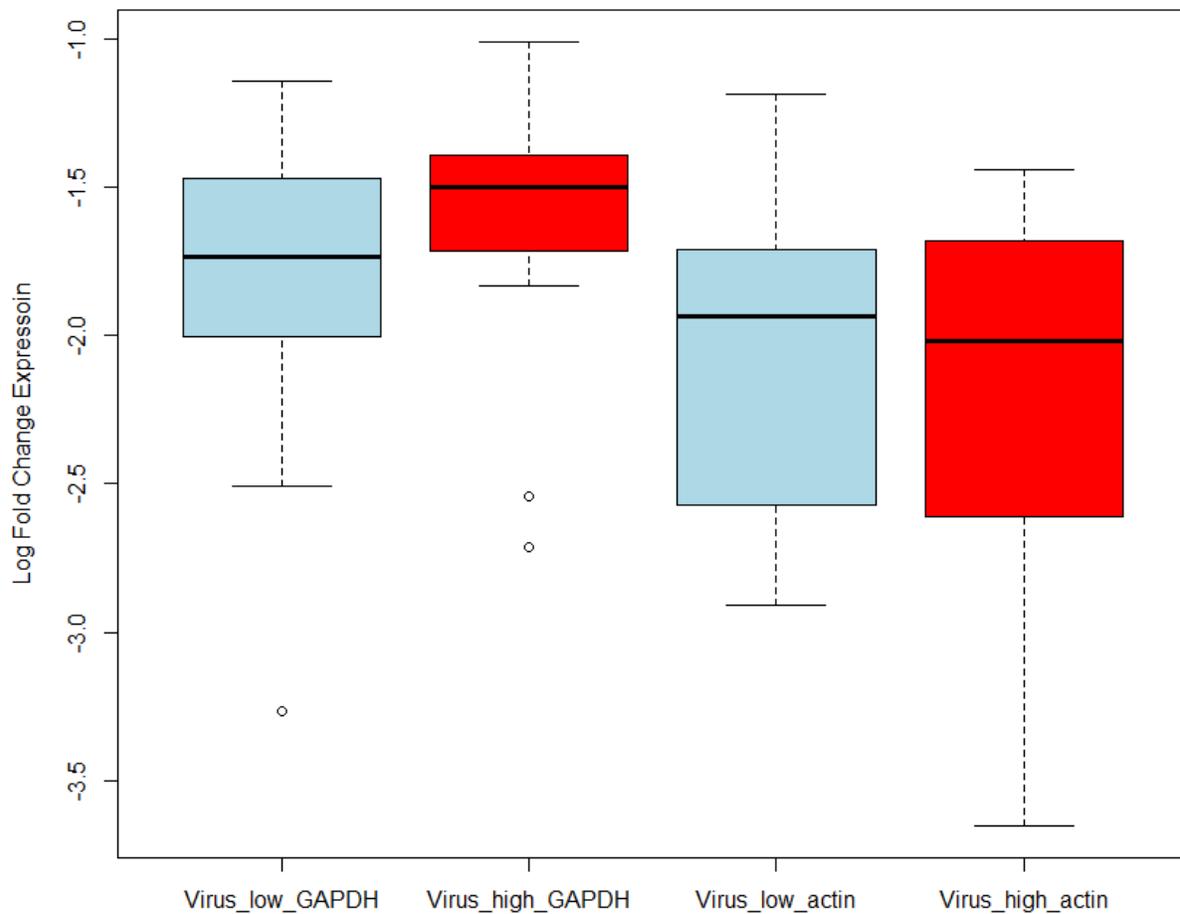


Figure 13. Log-transformed relative expression data of 28S 2070-2130 rRNA normalized against two reference genes, GAPDH and  $\beta$ -actin, between virus-low and virus-high bees ( $n = 31$ ) ( $p = 0.241$  for GAPDH;  $p = 0.678$  for  $\beta$ -actin:). No association between this polyA region of 28S rRNA and viral infection was established. The differences in rRNA expression values solely based on the reference gene suggest that the conditions of the experiment may have manipulated the expression of the reference gene.

# References

- Aizen, M., Garibaldi, L., Cunningham, S., & Klein, A. M. (2008). Long-term global trends in crop yield and production reveal no current pollination shortage but increasing pollinator dependency. *Current Biology*, *18*(20), 1572-5. doi:10.1016/j.cub.2008.08.066
- Alaux, C., Sinha, S., Hasadsri, L., Hunt, G. J., Guzmán-Novoa, E., DeGrandi-Hoffman, G., Uribe-Rubio, J. L., Southey, B. R., Rodriguez-Zas, S. & Robinson, G.E. (2009). Honey bee aggression supports a link between gene regulation and behavioral evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(36), 15400-5. doi:10.1073/pnas.0907043106
- Albert, Š., Gätschenberger, H., Azzami, K., Gimple, O., Grimmer, G., Sumner, S., Fujiyuki, T., Tautz, J. & Mueller, M. (2011). Evidence of a novel immune responsive protein in the Hymenoptera. *Insect Biochemistry and Molecular Biology*, *41*(12), 968-81. doi:10.1016/j.ibmb.2011.09.006
- Amdam, G. , Hartfelder, K., Norberg, K., Hagen, A., Omholt, S. W., & Stig, W. (2004). Altered physiology in worker honey bees (Hymenoptera: Apidae) infested with the mite *Varroa destructor* (Acari: Varroidae): a factor in colony loss during overwintering? *Journal of Economic Entomology*, *97*(3), 741-747. doi:10.1603/0022-0493(2004)097[0741:APIWHB]2.0.CO;2
- Anderson, D. L., & Trueman, J. W. H. (2000). *Varroa jacobsoni* (Acari: Varroidae) is more than one species. *Experimental and Applied Acarology*, *24*, 165-189.
- Antúñez, K., Martín-Hernández, R., Prieto, L., Meana, A., Zunino, P., & Higes, M. (2009). Immune suppression in the honey bee (*Apis mellifera*) following infection by *Nosema ceranae* (Microsporidia). *Environmental Microbiology*, *11*(9), 2284-90. doi:10.1111/j.1462-2920.2009.01953.x
- Arjona, A., Wang, P., Montgomery, R. R., & Fikrig, E. (2011). Innate immune control of West Nile virus infection. *Cellular Microbiology*, *13*(11), 1648-58. doi:10.1111/j.1462-5822.2011.01649.x
- Aubert, M. (2008). Impact of virus infection in honey bees. In A. M, B. V. Ball, I. Fries, M. R.F.A., N. Milani, & I. Bernardinelli (Eds.), *Virology and the Honey Bee* (pp. 233–253). Luxembourg: European Communities.

- Avadhanula, V., Weasner, B. P., Hardy, G. G., Kumar, J. P., & Hardy, R. W. (2009). A novel system for the launch of alphavirus RNA synthesis reveals a role for the Imd pathway in arthropod antiviral response. *PLoS Pathogens*, 5(9), e1000582. doi:10.1371/journal.ppat.1000582
- Azzami, K., Ritter, W., Tautz, J., & Beier, H. (2012). Infection of honey bees with acute bee paralysis virus does not trigger humoral or cellular immune responses. *Archives of Virology*, 157(4), 689-702. doi:10.1007/s00705-012-1223-0
- Bailey, L., Ball, B. V., & Perry, J. N. (1983). Association of viruses with two protozoal pathogens of the honey bee. *Annals of Applied Biology*, 103(1), 13-20. doi:10.1111/j.1744-7348.1983.tb02735.x
- Bailey, L., & Woods, R. D. (1977). Two more small RNA viruses from honey bees and further observations on Sacbrood and Acute Bee Paralysis viruses. *J Gen Viro*, 37(1), 175-182. doi:10.1099/0022-1317-37-1-175
- Ball, B. V., & Allen, M. F. (1988). The prevalence of pathogens in honey bee (*Apis mellifera*) colonies infested with the parasitic mite *Varroa jacobsoni*. *Annals of Applied Biology*, 113, 237-244.
- Boncrisiani, H., Di Prisco, G., Pettis, J., Hamilton, M., & Chen, Y. (2009). Molecular approaches to the analysis of deformed wing virus replication and pathogenesis in the honey bee, *Apis mellifera*. *Virology Journal*, 6(1), 221.
- Boncrisiani, H. F., Di Prisco, G., Pettis, J. S., Hamilton, M., & Chen, Y. P. (2009). Molecular approaches to the analysis of deformed wing virus replication and pathogenesis in the honey bee, *Apis mellifera*. *Virology Journal*, 6(221). doi:10.1186/1743-422X-6-221
- Bonderoff, J. M., Larey, J. L., & Lloyd, R. E. (2008). Cleavage of poly(A)-binding protein by poliovirus 3C proteinase inhibits viral internal ribosome entry site-mediated translation. *Journal of Virology*, 82(19), 9389-99. doi:10.1128/JVI.00006-08
- Bowen-Walker, P., Martin, S., & Gunn, a. (1999). The transmission of deformed wing virus between honeybees (*Apis mellifera* L.) by the ectoparasitic mite *Varroa jacobsoni* Oud. *Journal of Invertebrate Pathology*, 73(1), 101-6. doi:10.1006/jipa.1998.4807
- Bromenshenk, J. J., Henderson, C. B., Wick, C. H., Stanford, M. F., Zulich, A. W., Jabbour, R. E., Deshpande, S. V., McCubbin, P.E., Seccomb, R.A., Welch, P.M., Williams, T., Firth, D.R., Skowronski, E., Lehmann, M.M., Bilimoria, S.L., Gress, J., Wanner, K.W. & Cramer, R.A. (2010). Iridovirus and microsporidian linked to honey bee colony decline. *PLoS One*, 5(10), e13181.

- Calderone, N. W. (2012). Insect Pollinated Crops, Insect Pollinators and US Agriculture: Trend Analysis of Aggregate Data for the Period 1992–2009. (G. Smagghe, Ed.) *PLoS One*, 7(5), e37235. doi:10.1371/journal.pone.0037235
- Carter, M., & Genersch, E. (2008). Molecular characterisation of honey bee viruses. In A. M, B. V. Ball, I. Fries, M. R.F.A., N. Milani, & I. Bernardinelli (Eds.), *Virology and the Honey Bee* (pp. 85-120). Luxembourg: European Communities.
- Chaimanee, V., Chantawannakul, P., Chen, Y., Evans, J. D., & Pettis, J. S. (2012). Differential expression of immune genes of adult honey bee (*Apis mellifera*) after inoculated by *Nosema ceranae*. *Journal of Insect Physiology*, (May). doi:10.1016/j.jinsphys.2012.04.016
- Chantawannakul, P., Ward, L., Boonham, N., & Brown, M. (2006). A scientific note on the detection of honeybee viruses using real-time PCR (TaqMan) in *Varroa* mites collected from a Thai honeybee (*Apis mellifera*) apiary. *Journal of Invertebrate Pathology*, 91(1), 69-73. doi:10.1016/j.jip.2005.11.001
- Chauzat, M., Faucon, J., Martel, A., Lachaize, J., Cougoule, N. & Aubert, M. (2006). A survey of pesticide residues in pollen loads collected by honey bees in France. *Apiculture and Social Insects*, 99(2), 253-262.
- Chen, Y.P., & Evans, J. (2007). Historical presence of Israel acute paralysis virus in the United States. *American Bee Journal*, (December 2007), 1027-8.
- Chen, Y.P., Evans, J. D., Smith, I. B., & Pettis, J. S. (2008). *Nosema ceranae* is a long-present and wide-spread microsporidian infection of the European honey bee (*Apis mellifera*) in the United States. *Journal of Invertebrate Pathology*, 97(2), 186-8. doi:10.1016/j.jip.2007.07.010
- Chen, Y.P., Evans, J., & Feldlaufer, M. (2006). Horizontal and vertical transmission of viruses in the honey bee, *Apis mellifera*. *Journal of Invertebrate Pathology*, 92(3), 152-9. doi:10.1016/j.jip.2006.03.010
- Chen, Y P, Higgins, J. A., & Feldlaufer, M. F. (2005). Quantitative real-time reverse transcription-PCR analysis of Deformed Wing Virus infection in the honeybee (*Apis mellifera* L.). *Environmental Microbiology*, 71(1), 436-441. doi:10.1128/AEM.71.1.436
- Chen, Y P, Pettis, J. S., Collins, A., & Feldlaufer, M. F. (2006). Prevalence and transmission of honeybee viruses. *Applied and Environmental Microbiology*, 72(1), 606-611. doi:10.1128/AEM.72.1.606
- Chen, Y.P., & Siede, R. (2007). Honey bee viruses. *Advances in Virus Research*, 70(07), 33-80. doi:10.1016/S0065-3527(07)70002-7

- Chen, Y.P., Zhao, Y., Hammond, J., Hsu, H., Evans, J., & Feldlaufer, M. (2004). Multiple virus infections in the honey bee and genome divergence of honey bee viruses. *Journal of Invertebrate Pathology*, 87(2-3), 84-93. doi:10.1016/j.jip.2004.07.005
- Clarke, T. E., & Clem, R. J. (2003). Insect defenses against virus infection: the role of apoptosis. *International Reviews of Immunology*, 22(5/6), 401-424.
- Claudianos, C., Ranson, H., Johnson, R. M., Biswas, S., Schuler, M. A., Berenbaum, M. R., Feyereisen, R. & Oakeshott, J.G. (2006). A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee. *Insect Molecular Biology*, 15(5), 615-636.
- Cocquet, J., Chong, A., Zhang, G., & Veitia, R. (2006). Reverse transcriptase template switching and false alternative transcripts. *Genomics*, 88(1), 127-31. doi:10.1016/j.ygeno.2005.12.013
- Core, A., Runckel, C., Ivers, J., Quock, C., Siapno, T., Denault, S., Brown, B., et al. (2012). A new threat to honey bees, the parasitic phorid fly *Apocephalus borealis*. *PLoS One*, 7(1), e29639. doi:10.1371/journal.pone.0029639
- Costa, C., Tanner, G., Lodesani, M., Maistrello, L., & Neumann, P. (2011). Negative correlation between *Nosema ceranae* spore loads and deformed wing virus infection levels in adult honey bee workers. *Journal of Invertebrate Pathology*, 108(3), 224-5. doi:10.1016/j.jip.2011.08.012
- Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N. A., Quan, P. L., Briese, T., Hornig, M. Geiser, D.M., Martinson, V., vanEngelsdorp, D., Kalkstein, A.L, Drysdale, A., Hui, J., Zhai, J., Cui, L., Hutchinson, S.K., Smons, J.F., Egholm, M., Pettis, J.S. & Lipkin, W.I. (2007). A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, 318, 283-287.
- Dainat, B., Evans, J. D., Chen, Y. P., Gauthier, L., & Neumann, P. (2012). Predictive markers of honey bee colony collapse. *PLoS One*, 7(2), e32151.
- Dainat, B., Evans, J. D., Chen, Y. P., & Neumann, P. (2011). Sampling and RNA quality for diagnosis of honey bee viruses using quantitative PCR. *Journal of Virological Methods*, 174(1-2), 150-152. doi:10.1016/j.jviromet.2011.03.029
- Dainat, B., vanEngelsdorp, D., & Neumann, P. (2012). Colony collapse disorder in Europe. *Environmental Microbiology Reports*, 4(1), 123-125. doi:10.1111/j.1758-2229.2011.00312.x

- De Gregorio, E., Spellman, P. T., Tzou, P., Rubin, G. M., & Lemaitre, B. (2002). The Toll and Imd pathways are the major regulators of the immune response in *Drosophila*. *The EMBO Journal*, 21(11), 2568-79. doi:10.1093/emboj/21.11.2568
- Dostert, C., Jouanguy, E., Irving, P., Troxler, L., Galiana-Arnoux, D., Hetru, C., Hoffmann, J. a, et al. (2005). The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of *Drosophila*. *Nature Immunology*, 6(9), 946-53. doi:10.1038/ni1237
- Draghici, S., Khatri, P., Eklund, A. C., & Szallasi, Z. (2006). Reliability and reproducibility issues in DNA microarray measurements. *Trends in Genetics : TIG*, 22(2), 101-9. doi:10.1016/j.tig.2005.12.005
- Drake, J. W., Charlesworth, B., Charlesworth, D., & Crow, J. F. (1998). Rates of spontaneous mutation. *Genetics*, 148(4), 1667-86.
- Drummond, A., Ashton, B., Buxton, S., Cheung, M., Cooper, A., Duran, C., Field, M., et al. (2011). Geneious Pro. Retrieved from <http://www.geneious.com/>
- Ellis, J. D., Evans, J. D., & Pettis, J. (2010). Colony losses, managed colony population decline, and Colony Collapse Disorder in the United States. *Journal of Apicultural Research*, 49(1), 134-136. doi:10.3896/IBRA.1.49.1.30
- Evans, J. D., Aronstein, K., Chen, Y. P., Hetru, C., Imler, J.-L., Jiang, H., Kanost, M., et al. (2006). Immune pathways and defence mechanisms in honey bees *Apis mellifera*. *Insect Molecular Biology*, 15(5), 645-656. Blackwell Publishing Ltd. doi:10.1111/j.1365-2583.2006.00682.x
- Fernández-Miragall, O., & Martínez-Salas, E. (2007). In vivo footprint of a picornavirus internal ribosome entry site reveals differences in accessibility to specific RNA structural elements. *The Journal of General Virology*, 88(11), 3053-62. doi:10.1099/vir.0.83218-0
- Fievet, J., Tentcheva, D., Gauthier, L., de Miranda, J., Cousserans, F., Colin, M., & Bergoin, M. (2006). Localization of deformed wing virus infection in queen and drone *Apis mellifera* L. *Virology Journal*, 3(1), 16.
- Foster, L. J. (2011). Interpretation of data underlying the link between colony collapse disorder (CCD) and an invertebrate iridescent virus. *Molecular & Cellular Proteomics : MCP*, 10(3), M110.006387. doi:10.1074/mcp.M110.006387
- Fujiyuki, T., Ohka, S., Takeuchi, H., Ono, M., Nomoto, A., & Kubo, T. (2006). Prevalence and phylogeny of Kakugo virus, a novel insect picorna-like virus that infects the honeybee (*Apis mellifera* L.), under various colony conditions. *Journal of Virology*, 80(23), 11528-38. doi:10.1128/JVI.00754-06

- Gallai, N., Salles, J. M., Settele, J., & Vaissiere, B. E. (2009). Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. *Ecological Economics*, 68, 810-821.
- Gauthier, L., Tentcheva, D., Tournaire, M., Dainat, B., Cousserans, F., Colin, M., & Bergoin, M. (2007). Viral load estimation in asymptomatic honey bee colonies using the quantitative RT-PCR technique. *Apidologie* 38(5), 426-435. doi:10.1051/apido:2007026
- Genersch, E., Yue, C., Fries, I., & de Miranda, J. R. (2006). Detection of Deformed wing virus, a honey bee viral pathogen, in bumble bees (*Bombus terrestris* and *Bombus pascuorum*) with wing deformities. *Journal of Invertebrate Pathology*, 91(1), 61-3. doi:10.1016/j.jip.2005.10.002
- Gilboa, E., Mitra, S. W., Goof, S., & Baltimore, D. (1979). A Detailed Model of Reverse Transcription of Crucial Aspects and Tests. *Cell*, 18(1), 93-100.
- Gillespie, J. J., Johnston, J. S., Cannone, J. J., & Gutell, R. R. (2006). Characteristics of the nuclear (18S, 5.8S, 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta : Hymenoptera): structure , organization , and retrotransposable elements. *Insect molecular biology*, 15(5), 657-686.
- Gisder, S., Aumeier, P., & Genersch, E. (2009). Deformed wing virus: replication and viral load in mites (*Varroa destructor*). *Journal of General Virology*, 90(2), 463-467. doi:10.1099/vir.0.005579-0
- Graveley, B. R., Brooks, A. N., Carlson, J. W., Duff, M. O., Landolin, J. M., Yang, L., Artieri, C. G., et al. (2011). The developmental transcriptome of *Drosophila melanogaster*. *Nature*, 471(7339), 473-9. Nature Publishing Group. doi:10.1038/nature09715
- Gregory, P. G., Evans, J. D., Rinderer, T. E., & De Gruzman, L. (2005). Conditional immune-gene suppression of honeybees parasitized by Varroa mites. *J Insect Scien*, 5, 1-5.
- Grzechnik, P., & Kufel, J. (2008). Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. *Molecular Cell*, 32(2), 247-258. doi:DOI: 10.1016/j.molcel.2008.10.003
- Habayeb, M. S., Ekengren, S. K., & Hultmark, D. (2006). Nora virus, a persistent virus in *Drosophila*, defines a new picorna-like virus family. *Journal of General Virology*, 87(10), 3045-3051. doi:10.1099/vir.0.81997-0
- Hannon, G. J. (2002). RNA interference. *Nature*, 418, 24-26.
- Hedges, L. M., & Johnson, K. N. (2008). Induction of host defence responses by *Drosophila C* virus. *The Journal of General Virology*, 89, 1497-1501. doi:10.1099/vir.0.83684-0

- Henry, M., Beguin, M., Requier, F., Rollin, O., Odoux, J.-F., Aupinel, P., Aptel, J., Aptel, J., Tchamitchian, S., Decourtye, A. (2012). A common pesticide decreases foraging success and survival in honey bees. *Science*, 336, 348-350. doi:10.1126/science.1215039
- Higes, M., Garcia-Palencia P., Martín-Hernández, R. & Meana, A. (2007). Experimental Infection of *Apis mellifera* honeybees with *Nosema ceranae* (Microsporidia). *Journal of Invertebrate pathology*, 94, 211-217
- Highfield, A. C., El Nagar, A., Mackinder, L. C., Noel, L. M., Hall, M. J., Martin, S. J., & Schroeder, D. C. (2009). Deformed wing virus implicated in overwintering honeybee colony losses. *Appl Environ Microbiol*, 75, 7212-7220.
- Hudson, M. E. (2008). Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular Ecology Resources*, 8(1), 3-17. doi:10.1111/j.1471-8286.2007.02019.x
- Hunter, W. (2010). Medium for development of bee cell cultures (*Apis mellifera*: Hymenoptera: Apidae). *In Vitro Cellular & Developmental Biology - Animal*, 46(2), 83-86. doi:10.1007/s11626-009-9246-x
- Johnson, R. M., Evans, J. D., Robinson, G. E., & Berenbaum, M. R. (2009). Changes in transcript abundance relating to colony collapse disorder in honey bees (*Apis mellifera*). *Proceedings of the National Academy of Sciences of the United States of America*, 106(35), 14790-5. doi:10.1073/pnas.0906970106
- Klee, J., Besana, A. M., Genersch, E., Gisder, S., Nanetti, A., Tam, D. Q., Chinh, T. X., et al. (2007). Widespread dispersal of the microsporidian *Nosema ceranae*, an emergent pathogen of the western honey bee, *Apis mellifera*. *Journal of Invertebrate Pathology*, 96(1), 1-10. doi:10.1016/j.jip.2007.02.014
- Kong, Q., Stockinger, M. P., Chang, Y., Tashiro, H., & Lin, C.-liang G. (2008). The presence of rRNA sequences in polyadenylated RNA and its potential functions. *Biotechnology Journal*, 3(8), 1041-1046. doi:10.1002/biot.200800122
- Lange, H., Holec, S., Cognat, V., Pieuchot, L., Le Ret, M., Canaday, J., & Gagliardi, D. (2008). Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in *Arabidopsis thaliana*. *Molecular and Cellular Biology*, 28(9), 3038-3044. doi:10.1128/MCB.02064-07
- Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology*, 10(3), R25.
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_T}$  Method. *Nature Methods*, 25(4), 402-408. doi:DOI: 10.1006/meth.2001.1262

- Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M., & Gilad, Y. (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research*, *18*(9), 1509-17. doi:10.1101/gr.079558.108
- Martin, S. J., Highfield, a. C., Brettell, L., Villalobos, E. M., Budge, G. E., Powell, M., Nikaido, S. & Schoreder, D.C. (2012). Global honey bee viral landscape altered by a parasitic mite. *Science*, *336*(6086), 1304-1306. doi:10.1126/science.1220941
- Martín-Hernández, R., Meana, A., Prieto, L., Martínez-Salvador, A., Garrido-Bailón, E. Higes, M. (2007). Outcome of colonization of *Apis mellifera* by *Nosema ceranae*. *Applied and Environmental Microbiology*, *73*, 6331-6338.
- Martínez-Salas, E., Ramos, R., Lafuente, E., & López de Quinto, S. (2001). Functional interactions in internal translation initiation directed by viral and cellular IRES elements. *The Journal of General Virology*, *82*(Pt 5), 973-84.
- Mauro, V. P., & Edelman, G. M. (1997). rRNA-like sequences occur in diverse primary transcripts: Implications for the control of gene expression. *Proceedings of the National Academy of Sciences*, *94*(2), 422-427. doi:10.1073/pnas.94.2.422
- Miranda, J. R. (2008). Diagnostic techniques for virus detection in honey bees. In M. Aubert, B. Ball, I. Fries, R. Moritz, N. Milani, & I. Bernardinelli (Eds.), *Virology and the Honey Bee* (pp. 121-232). Luxembourg: European Communities.
- Miranda, J. R., & Genersch, E. (2010). Deformed wing virus. *Journal of Invertebrate Pathology*, *103 Suppl* (2010), S48-61. doi:10.1016/j.jip.2009.06.012
- Moritz, R. F. A., & Evans, J. D. (2008). Honey bee genomics and breeding for resistance to virus infections. In M. Aubert, B. Ball, I. Fries, R. Moritz, N. Milani, & I. Bernardinelli (Eds.), *Virology and the Honey Bee* (pp. 347-370). Luxembourg: European Communities.
- Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature* *5*(7), 1-8. doi:10.1038/NMETH.1226
- Mullin, C. a, Frazier, M., Frazier, J. L., Ashcraft, S., Simonds, R., vanEngelsdorp, D., & Pettis, J. S. (2010). High levels of miticides and agrochemicals in North American apiaries: implications for honey bee health. *PloS One*, *5*(3), e9754. doi:10.1371/journal.pone.0009754
- Munoz-Torres, M. C., Reese, J. T., Childers, C. P., Bennett, A. K., Sundaram, J. P., Childs, K. L., Anzola, J. M., et al. (2011). Hymenoptera Genome Database: integrated community resources for insect species of the order Hymenoptera. *Nucleic Acids Research* *39*(Database issue), D658-62. doi:10.1093/nar/gkq1145

- Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., & Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320(5881), 1344-9. doi:10.1126/science.1158441
- Navajas, M., Migeon, A., Alaux, C., Martin-Magniette, M. L., Robinson, G. E., Evans, J. D., Cros-  
Arteil, S., et al. (2008). Differential gene expression of the honey bee *Apis mellifera*  
associated with *Varroa destructor* infection. *BMC Genomics*, 9(1), 301.
- Nazzi, F., Brown, S. P., Annoscia, D., Del Piccolo, F., Di Prisco, G., Varricchio, P., Della Vedova, G.,  
et al. (2012). Synergistic parasite pathogen interactions mediated by host immunity can  
drive the collapse of honeybee colonies. (D. S. Schneider, Ed.) *PLoS Pathogens*, 8(6),  
e1002735. doi:10.1371/journal.ppat.1002735
- Nordström, S. (2003). Distribution of deformed wing virus within honey bee (*Apis mellifera*)  
brood cells infested with the ectoparasitic mite *Varroa destructor*. *Experimental & Applied  
Acarology*, 29(3-4), 293-302.
- Nordström, S., Fries, I., Aarhus, A., Hansen, H., & Korpela, S. (1999). Virus infections in Nordic  
honey bee colonies with no, low or severe *Varroa jacobsoni* infestations. *Apidologie*, 30(6),  
475-484.
- Oldroyd, B. P. (2007). What's killing American honey bees? *PLoS Biol*, 5(6), e168.
- Ongus, J. R., Peters, D., Bonmatin, J.-M., Bengsch, E., Vlak, J. M., & van Oers, M. M. (2004).  
Complete sequence of a picorna-like virus of the genus Iflavirus replicating in the mite  
*Varroa destructor*. *The Journal of General Virology*, 85(Pt 12), 3747-55.  
doi:10.1099/vir.0.80470-0
- Pestova, T. V., & Hellen, C. U. T. (2006). Translation, interrupted. *Nature Structural & Molecular  
Biology*, 13(2), 98-9. doi:10.1038/nsmb0206-98
- Pettis, J. S. (2004). A scientific note on *Varroa destructor* resistance to coumaphos in the United  
States. *Apidologie*, 35, 91-92. doi:10.1051/apido
- Ribière, M., Ball, B. V., & Aubert, M. (2008). Natural history and geographical distribution of  
honey bee viruses. In A. M, B. V. Ball, I. Fries, M. R.F.A., N. Milani, & I. Bernardinelli (Eds.),  
*Virology and the Honey Bee* (pp. 15-84). European Communities.
- Rosenkranz, P., Aumeier, P., & Ziegelmann, B. (2010). Biology and control of *Varroa destructor*.  
*J Invertebr Pathol*, 103(Supplement), S96-119.
- Rozen, S., & Skalesky, H. J. (2000). Primer3 on the WWW for general users and for biologist  
programmers. In S. Krawetz & S. Misener (Eds.), *Bioinformatics Methods and Protocols:  
Methods in Molecular Biology* (pp. 365-386). Totowa, NJ: Humana Press.

- Runckel, C., Flenniken, M. L., Engel, J. C., Ruby, J. G., Ganem, D., Andino, R., & DeRisi, J. L. (2011). Temporal analysis of the honey bee microbiome reveals four novel viruses and seasonal prevalence of known viruses, *Nosema*, and *Crithidia*. *PLoS One*, *6*(6), e20656.
- Scharlaken, B., de Graff, D., Goossens, K., Brunain, M., Peelman, L., & Jacobs, F. (2008). Reference gene selection for insect expression studies using quantitative real-time PCR: The head of the honeybee, *Apis mellifera*, after a bacterial challenge. *Journal of Insect Science*, *8*(33), 10pp.
- Schena, M., Shalon, D., Davis, R. W., & Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* *270*(5235), 467-70.
- Schmid-Hempel, P. (2005). Evolutionary ecology of insect immune defenses. *Annual Review of Entomology*, *50*(22), 529-51. doi:10.1146/annurev.ento.50.071803.130420
- Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nat. Protocols*, *3*(6), 1101-1108.
- Shen, M., Yang, X., Cox-Foster, D., & Cui, L. (2005). The role of *Varroa* mites in infections of Kashmir bee virus (KBV) and deformed wing virus (DWV) in honey bees. *Virology*, *342*(1), 141-149. doi:10.1016/j.virol.2005.07.012
- Slomovic, S., Laufer, D., Geiger, D., & Schuster, G. (2006a). Polyadenylation of ribosomal RNA in human cells. *Nucleic Acids Research*, *34*(10), 2966-2975. doi:10.1093/nar/gkl357
- Slomovic, S., Portnoy, V., Liveanu, V., & Schuster, G. (2006b). RNA polyadenylation in prokaryotes and organelles; different tails tell different tales. *Critical Reviews in Plant Sciences*, *25*(1), 65-77.
- Spivak, M., & Reuter, G. S. (2001). *Varroa destructor* infestation in untreated honey bee (Hymenoptera: Apidae) colonies selected for hygienic behavior. *Journal of Economic Entomology*, *94*(2), 326-31.
- Strand, M. R. (2008). The insect cellular immune response. *Insect Science*, *15*(1), 1-14. doi:10.1111/j.1744-7917.2008.00183.x
- Sumpter, D. J. T., & Martin, S. J. (2004). The dynamics of virus epidemics in *Varroa*-infested honey bee colonies. *Journal of Animal Ecology*, *73*(1), 51-63. doi:10.1111/j.1365-2656.2004.00776.x
- Tautz, D., Hancock, J., Webb, D., Tautz, C., & Dover, G. (1988). Complete sequences of the rRNA genes of *Drosophila melanogaster*. *Mol. Biol. Evol.*, *5*(4), 366-376.

- Tentcheva, D., Gauthier, L., Zappulla, N., Dainat, B., Cousserans, F., Colin, M. E., & Bergoin, M. (2004). Prevalence and seasonal variations of six bee viruses in *Apis mellifera* L. and *Varroa destructor* mite populations in France. *Applied and Environmental Microbiology*, 70(12), 7185-7191. doi:10.1128/AEM.70.12.7185-7191.2004
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., & Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology*, 3(7), RESEARCH0034.1-0037.11.
- vanEngelsdorp, D., Evans, J. D., Saegerman, C., Mullin, C., Haubruge, E., Nguyen, B. K., Frazier, M., et al. (2009). Colony collapse disorder: a descriptive study. *PLoS One*, 4(8), e6481.
- vanEngelsdorp, D., Hayes Jr., J., Underwood, R. M., Caron, D., & Pettis, J. (2011). A survey of managed honey bee colony losses in the USA, fall 2009 to winter 2010. *Journal of Apicultural Research*, 50(1), 1-10. doi:10.3896/IBRA.1.50.1.01
- Wang, L., Feng, Z., Wang, X., Wang, X., & Zhang, X. (2010). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics (Oxford, England)*, 26(1), 136-8. doi:10.1093/bioinformatics/btp612
- Wang, X.-H., Aliyari, R., Li, W.-X., Li, H.-W., Kim, K., Carthew, R., Atkinson, P., et al. (2006). RNA interference directs innate immunity against viruses in adult *Drosophila*. *Science (New York, N.Y.)*, 312(5772), 452-4. doi:10.1126/science.1125694
- Washburn, J. O., Kirkpatrick, B. A., & Volkman, L. E. (1996). Insect protection against viruses. *Nature*, 383, 767.
- Weinstein, J. N., Scherf, U., Lee, J. K., Nishizuka, S., Gwadry, F., Bussey, A. K., Kim, S., et al. (2002). The bioinformatics of microarray gene expression profiling. *Cytometry*, 47(1), 46-49. doi:10.1002/cyto.10041
- Winnebeck, E. C., Millar, C. D., & Warman, G. R. (2010). Why does insect RNA look degraded? *Journal of Insect Science*, 10(159). doi:insectscience.org/10.159
- Winston, M. L. (1991). *The Biology of the Honey Bee*. Cambridge: Harvard University Press.
- Wold, B., & Myers, R. M. (2008). Sequence census methods for functional genomics. *Nature Methods*, 5(1), 19-21. doi:10.1038/NMETH1157
- Yang, X. Y., & Cox-Foster, D. L. (2005). Impact of an ectoparasite on the immunity and pathology of an invertebrate: evidence for host immunosuppression and viral amplification. *Proc Natl Acad Sci USA*, 102(11), 7470-7475.

- Yue, C., & Genersch, E. (2005). RT-PCR analysis of Deformed wing virus in honeybees (*Apis mellifera*) and mites (*Varroa destructor*). *The Journal of General Virology*, *86*(12), 3419-24. doi:10.1099/vir.0.81401-0
- Zambon, R. a, Nandakumar, M., Vakharia, V. N., & Wu, L. P. (2005). The Toll pathway is important for an antiviral response in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(20), 7257-62. doi:10.1073/pnas.0409181102
- Zhang, Z., Schwartz, S., Wagner, L., & Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J Comput Bio* *7*(1-2), 203-214. doi:10.1089/10665270050081478.
- Zhang, Z. (2000). Notes on *Varroa destructor* (Acari : Varroidae ) parasitic on honeybees in New Zealand. *Systematic & Applied Acarology Special Publications*, *5*, 9-14.

# Appendix A: Shell script to find polyA sequences

```
perl full_and_partial_adap_trim.pl $FILE #trim adapters
perl convert_FQ_format.pl fq2fa $FILE.txt_adap_trimmed > $FILE.fa
#convert fastq to fasta
perl count_with_headers.pl $FILE.fa > $FILE.derep #dereplicates reads

for FILE in *derep
do
  `perl -i -pe 's/\n/\t/g' $FILE` #replaces new lines with tabs.
  makes dereplicated file into a single line
  `perl -i -pe 's/\t\>/\n\>/g' $FILE` #puts header and sequence
  data in a single line
  wc -l $FILE
  egrep -v "A{71}" $FILE | egrep "A{10}$" > $FILE.as #throws away
  junk sequence. finds 10as at the end
  wc -l $FILE.as
  egrep -v "A{71}" $FILE | egrep "T{10}$" > $FILE.ts #throws away
  junk sequence. finds 10ts at the end
  wc -l $FILE.ts
  egrep -v "A{71}" $FILE | egrep "[[:space:]]A{10}" > $FILE.startas
  #throws away junk sequence. finds 10as at the beginning
  wc -l $FILE.startas
  egrep -v "A{71}" $FILE | egrep "[[:space:]]T{10}" > $FILE.startts
  #throws away junk sequence. finds 10ts at the beginning
  wc -l $FILE.startts
  cat $FILE.as $FILE.ts $FILE.startas $FILE.startts >> $FILE.10
  #combine into one file for bowtie
  uniq $FILE.10 | wc -l
  `perl -i -pe 's/A{10,70}$/\t/g' $FILE.10` #throw away all
  polyadenylated sequences for better matching with bowtie
  `perl -i -pe 's/T{10,70}$/\t/g' $FILE.10`
  `perl -i -pe 's/[[:space:]]A{10,70}/\t/g' $FILE.10`
  `perl -i -pe 's/[[:space:]]T{10,70}/\t/g' $FILE.10`
  awk '{print$1"_"$2"\n"$3}' $FILE.10 > $FILE.10.fa #reformat to
  fasta
  head $FILE.10.fa
  bowtie -f rRNA_Apis_melliferaT $FILE.10.fa > $FILE.bowtieoutput
  grep -c "18S" $FILE.bowtieoutput
  grep "18S" $FILE.bowtieoutput > $FILE.18S
  awk '{print">"$1"_"$3"\n"$5}' $FILE.18S > $FILE.18S.fa
  grep -c "28S" $FILE.bowtieoutput
  grep "28S" $FILE.bowtieoutput > $FILE.28S.fa
  awk '{print">"$1"_"$3"\n"$5}' $FILE.28S > $FILE.28S.fa
```

```
grep -c "ITS1" $FILE.bowtieoutput
grep -c "ITS2" $FILE.bowtieoutput
grep -c "5S" $FILE.bowtieoutput
grep -c "12S" $FILE.bowtieoutput
grep -c "16S" $FILE.bowtieoutput
done
```

# Appendix B: Perl script to find SNPs

```
#!/usr/bin/perl -w
# This script checks for the SNPs present in clustal alignment of DNA
seq and lists the counts.
```

```
my $outputdata;
open IN, "$ARGV[0]" or die "can't open input file\n";
open OUT, ">$ARGV[0].SNPs" or die "can't open output file";
my %base_info;
my $id;
while (<IN>){

    if(/(^S{10})(S+)/){
        $base_info{$1}.$2;
        $id=$1;
    }
}

my $total_sequence;
foreach my $variable (keys %base_info){
    #print "$variable\t$base_info{$variable}\n ";
    $total_sequence++;
}

my $total_positions = length($base_info{$id});
print "Length of alignment: $total_positions\nTotal number of seq in
the alignment: $total_sequence\nOutput stored as file:
$ARGV[0].SNPs\n";

#my $runningcount = 1;

print OUT "Pos\tA\tT\tG\tC\t-\n";
for ($n=0; $n<$total_positions; $n++){
    my $key_count=0;
    my %count_hash = (
        "A"=>"0",
        "T"=>"0",
        "G"=>"0",
        "C"=>"0",
        "-"=>"0"
    );
    foreach my $seqname (keys %base_info){
        $nucleotide = substr($base_info{$seqname}, $n, 1);
        $count_hash{$nucleotide}++;
    }
}
```

```

    foreach my $nuc (keys %count_hash){
        if ($count_hash{$nuc}>0){
            unless ($nuc eq "-") {
                $key_count++;
            }
        }
    }
    if ($key_count >1){
        my $pos=$n+1;
        print OUT
"$pos\t$count_hash{A}\t$count_hash{T}\t$count_hash{G}\t$count_hash{C}\t
$count_hash{'-'}\n";
    }

}

close IN;
close OUT;

```